

STIC-Biotech/ChemLib

156005

NRFO

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 4:16 PM
To: STIC-Biotech/ChemLib
Subject: 10/735,419

For 10/735,419, pls search and interference search:

SID 10: against the NT and AA data bases.

THANKS!!

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

No art

(724,229 ID #10)

Rev. Swope

10-294aa

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:57:41 ; Search time 174 Seconds
(without alignments)
865.237 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNINALVCGNGPSLKNID.....IRLPSPDIKHYLKEKYANKNR 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1571	100.0	294	2	Q7BP25	Q7bp25 campylobact
2	823	52.4	291	2	Q9LAK3	Q9lak3 campylobact
3	816	51.9	291	2	Q938X6	Q938x6 campylobact
4	813	51.8	291	2	Q93CZ5	Q93cz5 campylobact
5	812	51.7	291	2	Q93MQ0	Q93mq0 campylobact
6	804	51.2	291	2	Q93D05	Q93d05 campylobact
7	804	51.2	291	2	Q9F0M9	Q9f0m9 campylobact
8	804	51.2	291	2	Q9L905	Q9l9q5 campylobact
9	680.5	43.3	430	2	Q9RGF1	Q9rgf1 campylobact
10	673.5	42.9	303	2	Q9CLP3	Q9clp3 pasteurella
11	441.5	28.1	231	1	Y352 HABIN	P24324 haemophilus
12	143	9.1	753	2	Q98S90	Q98s90 guillardia
13	137.5	8.8	2232	2	Q7RK88	Q7rk88 plasmodium
14	137.5	8.8	3869	2	Q7RNU4	Q7rn84 plasmodium
15	136	8.7	754	2	Q7RG14	Q7rgi4 plasmodium
16	136	8.7	4095	2	Q6LFD1	Q6lfd1 plasmodium
17	135.5	8.6	1676	2	Q7RI39	Q7ri39 plasmodium
18	133.5	8.5	974	2	Q8I304	Q8i304 plasmodium
19	133.5	8.5	1725	2	Q8IKU2	Q8iku2 plasmodium
20	132.5	8.4	1993	2	Q8ILC6	Q8ilc6 plasmodium
21	132	8.4	1847	2	Q8IDM0	Q8idm0 plasmodium
22	131	8.3	3787	2	Q8I1Z6	Q8i1z6 plasmodium
23	130.5	8.3	2120	2	Q7RQ47	Q7rq47 plasmodium
24	130.5	8.3	2705	2	Q8I5I0	Q8i5i0 plasmodium
25	130	8.3	683	2	Q7RST7	Q7rst7 plasmodium
26	130	8.3	1679	2	Q8ILL0	Q8ill0 plasmodium
27	129.5	8.2	471	2	Q98RR3	Q98rr3 mycoplasma
28	129.5	8.2	850	2	Q6KID3	Q6kid3 mycoplasma
29	128.5	8.2	961	2	Q7RMS2	Q7rms2 plasmodium
30	128.5	8.2	2385	2	Q962I6	Q962i6 plasmodium
31	128.5	8.2	2400	2	Q8I3U7	Q8i3u7 plasmodium

RESULT 1

Q7BP25

ID Q7BP25 PRELIMINARY; PRT; 294 AA.

AC Q7BP25; Q9PNF4;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein Cj1140 (2,3-sialyl transferase) (Alpha-2,3-

DE sialyltransferase).

GN Name=cstIII; Synonyms=cst-III; OrderedLocusNames=Cj1140;

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;

OC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=197;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSC57360;

RA Query P., Ewing C.P., Moran A.P., Trust T.J.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MSC 57360;

RX MEDLINE=20536403; PubMed=11083778;

RX DOI=10.1128/IAI.68.12.6656-6662.2000;

RA Query P., Ewing C.P., Hickey T.E., Prendergast M.M., Moran A.P.;

RT "Sialylation of lipooligosaccharide cores affects immunogenicity and

RT serum resistance of Campylobacter jejuni."

RL Infect. Immun. 68:6656-6662(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 43429; and ATCC 43429;

RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;

RA Gilbert M., Karwasaki M.-F., Bernatchez S., Young N.M., Taboada E.,

RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;

RT "The genetic basis for the variation in the lipo-oligosaccharide of

RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated

RT ganglioside mimics in the core oligosaccharide."

RL J. Biol. Chem. 277:327-337(2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 11168;

RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;

RA Parkhill J., Wren B.W., Mungall K.L., Kettle J.M., Churcher C.M.,

RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,

RA Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,

RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrall B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

RT reveals hypervariable sequences."

RL Nature 403:665-668(2000).

DR EMBL; AF195055; AAG29922.1; -

DR EMBL; AF257460; AAG37020.1; -

DR EMBL; AF400047; AAK85419.1; -

DR EMBL; AY044156; AAK73183.1; -

```

DR EMBL; AL139077; CAB73395.1; -.
DR PIR; G81318; G81318.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Complete proteome; Glycosyltransferase; Hypothetical protein;
KW Transferase.
SQ SEQUENCE 294 AA; 35127 MW; 1729C9C35797D78F CRC64;

Query Match 100.0%; Score 1571; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 2.2e-101;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSMNALVCGGSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFE 60
DB 1 MSMNALVCGGSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFE 60

QY 61 QYTSKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHGHEIIKKLDKFFAYIKYNEI 120
DB 61 QYTSKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHGHEIIKKLDKFFA 120

QY 121 YIKNEYINRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
DB 121 YIKNEYINRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180

QY 181 GFKNQKPKFINHSMACDLOALDYLMKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKPKFINHSMACDLOALDYLMKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

QY 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKR 294
DB 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKR 294

RESULT 2
Q9LAK3 PRELIMINARY; PRT; 291 AA.
AC Q9LAK3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-2,3/8-sialyltransferase.
GN Name-cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OH4384;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brieson J.R., Karwaski M.F., Michniewicz J.,
RA Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT mhz (1)h and (13)c NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 700297;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF130984; AAF31771.1; -.
DR EMBL; AF216647; AAL36462.1; -.
DR PDB; 1RO7; X-ray; A/B/C/D=1-259.
DR PDB; 1RO8; X-ray; A/B=1-259.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34544 MW; AB0741D5FCEDAB6A CRC64;

```

```

Query Match 52.4%; Score 823; DB 2; Length 291;
Best Local Similarity 54.9%; Pred. No. 2e-49;
Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFEQYYSKK 67
DB 5 IISGNGPSLKEIDYSRLPNDFVRCNQFYFEDRYFVGKDVYVFNPFVFEQYYSKKH 64

QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHGHEIIKKLDKFFAYIKYNEI 127
DB 65 LIQNEEYETELMCSNYNQAHLENENFVTKTFDYDPDAHLGYDFKQLKDFNAYFPFHEI 124

QY 128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCTGFKNQKF 187
DB 125 YFNQRTSGVYMCATVALGYKSIYISGIDFYQNGSS-YAFDTKQKNLLKLPNPFKNDS 183

QY 188 KFINHSMACDLOALDYLMKRYDNYIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
DB 184 HYGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242

QY 244 ILIPDKYAQERYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
DB 243 ILIPSEA----YGFPSKNINFKIKENIYIKLIKDLIRLPSDIKHYFKGK 291

RESULT 3
Q938X6 PRELIMINARY; PRT; 291 AA.
AC Q938X6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3/alpha-2,8-sialyltransferase.
GN Name-cstII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43460;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AY044868; AAK96001.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34576 MW; 7BAE5F6021A56F08 CRC64;

Query Match 51.9%; Score 816; DB 2; Length 291;
Best Local Similarity 54.3%; Pred. No. 6.1e-49;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFEQYYSKK 67
DB 5 IISGNGPSLKEIDYSRLPNDFVRCNQFYFEDRYFVGKDVYVFNPFVFEQYYSKKH 64

QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHGHEIIKKLDKFFAYIKYNEI 127
DB 65 LIQNEEYETELMCSNFNQAHLENQNVKTFDYDPDAHLGYDFKQLKDFNAYFPFHEI 124

QY 128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCTGFKNQKF 187
DB 125 YFNQRTSGVYMCATVALGYKSIYISGIDFYQNGSS-YAFDTKQKNLLKLPNPFKNDS 183

QY 188 KFINHSMACDLOALDYLMKRYDNYIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
DB 184 KFINHSMACDLOALDYLMKRYDNYIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243

```

```

Db      184 HYGHSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy      244 ILIPDKYAQERYG-----KKSRLENHYKLIKDLIRLPSDIKHYLEK 288
Db      243 ILIPSSA----YGFKNINFKIKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 4
Q93CZ5 PRELIMINARY; PRT; 291 AA.
ID Q93CZ5
AC Q93CZ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bifunctional alpha-2,3/-2,8-sialyltransferase.
GN Name=csst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF401529; AAL06004.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34504 MW; 50162DEB8D81558F CRC64;

Query Match 51.8%; Score 813; DB 2; Length 291;
Best Local Similarity 54.3%; Pred. No. 9.9e-49;
Matches 159; Conservative 43; Mismatches 73; Indels 18; Gaps 6;

Qy      8 LVCNGPSLKNIDYKRLPKQDFVRCNQFPEDRYFVGDKVYFNFVFFPEOYTSKK 67
Db      5 IISGNPSLKEIDYSLRPNDVDFRCNQFPEDRYFVGDKVYFNFVFFPEOYTSKK 64
Qy      68 LIQNEEYNIENIVCSINLEYIDGFQVDNFELYSDFDAFLGHEIHKLDKDFAYIKYNEI 127
Db      65 LIQNEEYELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFPKQKFNAYFKPHEI 124
Qy      128 YNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFKNQKF 187
Db      125 YFNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFKNQKF 183
Qy      188 KFTNHSMACDLQALDYLKMYDVNIYSL--NS--DEYFKLAPDIGSDVLSKPKKYIND 243
Db      184 HYGHSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy      244 ILIPDKYAQERYG-----KKSRLENHYKLIKDLIRLPSDIKHYLEK 288
Db      243 ILIPSSA----YGFKNINFKIKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 5
Q93MQ0 PRELIMINARY; PRT; 291 AA.
ID Q93MQ0
AC Q93MQ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3/-alpha-2,8-sialyltransferase.
GN Name=csst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

```

```

OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43438;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF400048; AAK91725.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34572 MW; BD1BEDEFF6F521E6 CRC64;

Query Match 51.7%; Score 812; DB 2; Length 291;
Best Local Similarity 54.3%; Pred. No. 1.2e-48;
Matches 159; Conservative 46; Mismatches 70; Indels 18; Gaps 6;

Qy      8 LVCNGPSLKNIDYKRLPKQDFVRCNQFPEDRYFVGDKVYFNFVFFPEOYTSKK 67
Db      5 IISGNPSLKEIDYSLRPNDVDFRCNQFPEDRYFVGDKVYFNFVFFPEOYTSKK 64
Qy      68 LIQNEEYNIENIVCSINLEYIDGFQVDNFELYSDFDAFLGHEIHKLDKDFAYIKYNEI 127
Db      65 LIQNEEYETELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFPKQKFNAYFKPHEI 124
Qy      128 YNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFKNQKF 187
Db      125 YLNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFKNQKF 183
Qy      188 KFTNHSMACDLQALDYLKMYDVNIYSL--NS--DEYFKLAPDIGSDVLSKPKKYIND 243
Db      184 HYGHSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy      244 ILIPDKYAQERYG-----KKSRLENHYKLIKDLIRLPSDIKHYLEK 288
Db      243 ILIPSSA----YGFKNINFKIKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 6
Q93D05 PRELIMINARY; PRT; 291 AA.
ID Q93D05
AC Q93D05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3-sialyltransferase.
GN Name=csst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43456;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF401528; AAL05990.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34640 MW; F294A04ACBA82882 CRC64;

```

```
Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 67
Db 5 ILAGNPSLKEIDYSLRPNDPFRVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 64
Qy 68 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 127
Db 65 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 124
Qy 128 YNRQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNCKTGFKNQKF 187
Db 125 YFNQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNCKTGFKNQKF 183
Qy 188 KFINHSMACDLQALDYLMKRYDNIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIQEK--NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLENLHYKLIKDLRLPSDIKHLYKEK 288
Db 243 ILIPSSA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291

RESULT 7
ID Q9F0M9 PRELIMINARY; PRT; 291 AA.
AC Q9F0M9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2.3-sialyltransferase.
GN Namecst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43432;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF215659; AAG43979.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34553 MW; E41B594ACD7280F8 CRC64;

Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 67
Db 5 ILAGNPSLKEIDYSLRPNDPFRVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 64
Qy 68 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 127
Db 65 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 124
Qy 128 YNRQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNCKTGFKNQKF 187
Db 125 YFNQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNCKTGFKNQKF 183
Qy 188 KFINHSMACDLQALDYLMKRYDNIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIQEK--NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLENLHYKLIKDLRLPSDIKHLYKEK 288
Db 243 ILIPSSA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291
```

```
Db 243 ILIPSSA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291

RESULT 8
Q9L9Q5 PRELIMINARY; PRT; 291 AA.
ID Q9L9Q5;
AC Q9L9Q5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-2.3-sialyltransferase.
GN Namecst-II; Synonyms=cst;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43446;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brisson J.R., Karwaski M.F., Michniewicz J.,
RA Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT mhz (1)h and (13)c NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43446;
RA Gilbert M.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=81-176;
RX MEDLINE=21655167; PubMed=11796612; DOI=10.1128/IAI.70.2.787-793.2002;
RA Guerry P., Szymanski C.M., Prendergast M.M., Hickey T.B., Ewing C.P.,
RA Pattarini D.L., Moran A.P.;
RT "Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
RT affects ganglioside mimicry and invasiveness in vitro.";
RL Infect. Immun. 70:787-793(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=GB11;
RX PubMed=14742567; DOI=10.1128/IAI.72.2.1162-1165.2004;
RA Gilbert M., Godschalk P.C., Karwaski M.F., Ang C.W., Van Belkum A.,
RA Li J., Wakarchuk W.W., Endtz H.P.;
RT "Evidence for Acquisition of the Lipooligosaccharide Biosynthesis
RT Locus in Campylobacter jejuni GB11, a Strain Isolated from a Patient
RT with Guillain-Barre Syndrome, by Horizontal Exchange.";
RL Infect. Immun. 72:1162-1165(2004).
DR EMBL; AF167344; AAF34137.1; -.
DR EMBL; AF305571; AAL09368.1; -.
DR EMBL; AY422197; AAR82875.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34610 MW; 1BE4A04ACD728228 CRC64;

Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 67
Db 5 ILAGNPSLKEIDYSLRPNDPFRVRCNQFYFEDRYFGKDVKYVFNFPFFQYVTSKK 64
Qy 68 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 127
Db 65 LIQNEEYNIENIVCSITNLEIDGQFVDNFELVFSDAFLGHEIIKKLKDFFAYIKNEI 124
Qy 128 YNRQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNCKTGFKNQKF 187
```

```

Db      125 YFNQRTSGVMCAVAIAIKYKIYSGIDFYQNGSS-YAFDTKQENLLKLPDFFNDRS 183
QY      188 KFINHSMACDQALDYLMKRYDVIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
Db      184 HYIGHSKNTDKALEFLEKTYKIKLYCLPNSLLANFIELANLNSNFIQEK-NNYTKD 242
QY      244 ILIPDKYAQRYVG-----KKSRLKENLHYKLIKOLIRLPSDIKHYLKEK 288
Db      243 ILIPSSSEA----YKFSKNINFKIKIKENVYKLIKOLRLPSDINGHYFKG 291

RESULT 9
Q9RCF1 PRELIMINARY; PRT; 430 AA.
AC Q9RCF1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3-sialyltransferase.
GN Name-cst-I;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OH4384;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brisson J.R., Karwasaki M.F., Michniewicz J.,
RA Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT mhz (1)h and (13)c NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
DR EMBL; AF130466; AAF13495.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 430 AA; 50396 MW; 1D03B6797169425C CRC64;

Query Match 43.3%; Score 680.5; DB 2; Length 430;
Best Local Similarity 49.7%; Pred. No. 2.4e-39;
Matches 145; Conservative 49; Mismatches 75; Indels 23; Gaps 8;

QY 1 MSNII-NALVCGNGPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFV 59
Db 12 VSKNMQNIITAGNGPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFV 71
QY 60 EQYTSKKLIQNEBYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIILKLDKDF 119
Db 72 QQYHTAKQLIKNEYIKNIFCSTENLPFIESNDFLHQFYFPPDAKLGVEVIENLKEFY 131
QY 120 AYIKYNIYRQRTSGVMCAVAIAIKYKIYSGIDFYQNTNNLYAFDNNKKNLNC 179
Db 132 AYIKYNIYRQRTSGVMCAVAIAIKYKIYSGIDFYQNTNNLYAFDNNKKNLNC 179
QY 180 TGFKNQKFEIN-HSMACDQALDYLMKRYDVIYSLNSD-----EYFKLAPDIGSDPVLS 234
Db 191 PGIKD--FKPSNCHSKEYDEALKLKSIIYKVIYALCDDSIILANHPFUSINNNFTLE 248
QY 235 KPKPKYINDILIPDKYAQRYVGKSKRLKEN-----LHYK--LIKDL 274
Db 249 NKHNSINDILLTNTDPGVSVFY--KNQKADKNKIMLNFNLYLHSLKONLIKFL 298

RESULT 10
Q9CLP3 PRELIMINARY; PRT; 303 AA.
AC Q9CLP3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein PM1174.
GN OrderedLocusNames=PM1174;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006157; AAK03258.1; -.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 303 AA; 35676 MW; BA7B6C8C7563B921 CRC64;

Query Match 42.9%; Score 673.5; DB 2; Length 303;
Best Local Similarity 46.8%; Pred. No. 5.1e-39;
Matches 137; Conservative 47; Mismatches 94; Indels 15; Gaps 7;

QY 8 LVCGNGPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFEQYTSKK 67
Db 14 IVAGNGESLSQIDYRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFEQYTSKK 73
QY 68 LIQNEBYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHE-IILKLDKDFYIKYNE 126
Db 74 LKRNNEFYVDNVLSSFNHPTVD-LEKSKQIALFIDVINGEKYLSKLTAFDVIYRYKE 132
QY 127 IYNRQRTSGVMCAVAIAIKYKIYSGIDFYQNTNNLYAFDNNKKNLNCCTGPKNQK 186
Db 133 LYENQRITSGVMCAVAIAIKYKIYSGIDFYQNTNNLYAFDNNKKNLNCCTGPKNQK 192
QY 187 FKFINHSMACDQALDYLMKRYDVIYSLNS-----DEYFKLAPDIGSD----FVLSKPK 238
Db 193 TLFYSYHSKIDLEALSFLQOHYVNFYSISPSPLSKFPPI-PTVEDDCETTFVAPLK-B 250
QY 239 KYINDILIPDKYAQRYVGKSKRLKENLHYKLIKOLIRLPSDIKHYLKEK 288
Db 251 NYINDILIPDFVFEKLTIVSKSRFHSNLIIVLRILDLKLPALKHYLKEK 303

RESULT 11
Y352 HAEIN
ID Y352 HAEIN STANDARD; PRT; 231 AA.
AC F24324;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein HI0352 (ORF1).
GN OrderedLocusNames=HI0352;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM 7004 / Serotype B;
RX MEDLINE=92065797; PubMed=1956282;
RA Maskell D.J., Szabo M.J., Butler P.D., Williams A.E., Moxon E.R.;
RT "Molecular analysis of a complex locus from Haemophilus influenzae
RT involved in phase-variable lipopolysaccharide biosynthesis.";
RL Mol. Microbiol. 5:1013-1022(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

```



```

Db 1652 LSNLDGRKSEQKNV---NIFPKKGNFLVQRKYQREGANNIVTFAKEKKGKSDVSNIL 1708
Qy 64 TSKKLIONEYNIENI-----VCSTINLEYIDGQFVDNPFELYSDFAPLGH 109
Db 1709 STNKLNIHKNSNDSIDSYEKIEKVRSGIINTPSKRYI-----HREFLSDSPETN 1760
Qy 110 EIIKKLAK--DPPAVIKYNE---IYNQRITSGVVMCATAVAGYKSIYISIDIPYQDTNN 164
Db 1761 SFRKNKSSETIQIPKPRDRTIRNN-----SDHMGISYDDKFI-----EKNV 1805
Qy 165 LYAFDNNKQNLNCKTGFKN-----QKFKPIN-HSMACDLQALDYLKRYDNNIYLSND 218
Db 1806 INRVSEKSK-----FKSNLKKKSNYVDIRSIKSKSTLEYFVEK-----D 1847
Qy 219 EYFKLAPDIGSD-FVLSKKPKKYINDILIPDKVAQERYGKSRLENLHYKL--INKLI 275
Db 1848 YLKQWTTSTISDSYEIKDKTKYKKNKTIPKKNKKEKIEKKKCKKKEKKKKDIL 1907
Qy 276 RLPSDIKHLYKRYANKN 293
Db 1908 IYPOIKSYVKEK-INKN 1924

RESULT 14
Q7RN84 PRELIMINARY; PRT; 3869 AA.
AC Q7RN84;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein.
GN Name=PY01937;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73233;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAL01001318; EAA16213.1; -.
DR InterPro; IPR002917; MMR_HSR1.
DR Pfam; PF01926; MMR_HSR1; 1.
SQ SEQUENCE 3869 AA; 457791 MW; E0DF5A5603674201 CRC64;

Query Match 8.8%; Score 137.5; DB 2; Length 3869;
Best Local Similarity 22.5%; Pred. No. 1.4; Indels 127; Gaps 20;
Matches 73; Conservative 48; Mismatches 77;

Qy 27 QFDV-----FRCNPFYEDRYFVGKDVKYVFPFVEQYTSKLIQNEYN 75
Db 3436 QYDVLHGCLLYINPFPNNQL---DHEFFIAKEIPLVLF-----LIQNKQVS 3478
Qy 76 -----IENIVCSTINLEYIDGQFVDNPFELYSDFAPLGH 108
Db 3479 RDIKFGENVNRTNISNRCVCDVNYTELKNVCNCINIDIKNEFINNIE----- 3529
Qy 109 HEIIKKLKDPFAYIKYN-----EIVN--RQRTISGVVMCATAVAGYKSIYISIDIF 159

```

```

Db 3530 -EKIPKYNVYIEYNNNMKMYEIEENAEIEEKIS--YQYNNRYKYNKKVAYFKGFNIY 3586
Qy 160 QDT-----NNLYAFDNNKQNLNCKTGFKNQKFKFINHSMACDLQALDYLKRYDYN 211
Db 3587 EKENYEQIEKFNKYCIVTYKKKNILEKCNLYNN-----YTV-YNNGN 3627
Qy 212 IYSLNSDEYFKLAPDIGSDPVLSSKKPKKYINDILIPDKVAQERYGKSRLENLHYKLI 271
Db 3628 YLVNTRYE-----DV-LDF--EKNPETYVN--VEEPY---KKFEKISNLEENI--KIE 3670
Qy 272 KD---LIRLPSDIKHLYKRYANKN 293
Db 3671 NEGFCVLVSLKNE-NNEVK---GNKN 3691

RESULT 15
Q7RG14 PRELIMINARY; PRT; 754 AA.
AC Q7RG14;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Probable tRNA modification GTPase time.
GN Name=PY04363;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAL01001318; EAA16213.1; -.
DR InterPro; IPR002917; MMR_HSR1.
DR Pfam; PF01926; MMR_HSR1; 1.
SQ SEQUENCE 754 AA; 88639 MW; 2826844083775506 CRC64;

Query Match 8.7%; Score 136; DB 2; Length 754;
Best Local Similarity 26.0%; Pred. No. 0.3;
Matches 81; Conservative 35; Mismatches 85; Indels 110; Gaps 19;

Qy 17 KNIDYKRL--PKQFD-----VFRCNQFYED-RYFVGKDVK--YVFPNPFVEQYTSK 66
Db 211 KNIESRKLKYKIYDINSNDIIDNVVAYFKSPNSYTGEDVVEIYCHGNPPIV----- 262
Qy 67 KLIQNEEYNIENIVCSTINLEYIDGQFVDNPFELYSDFAPLGHIEIK---KLKDFPFIK 123
Db 263 KEIMNAIDHNNIMYIINDE-----YNNDDWYFNQIEQNIQEQIETIFDLNNFVIRSKKG 301
Qy 124 YNEIYNQRITSGVVMCATAVAGYKS---IYISGIDFYQ-----DTNN----- 164
Db 302 HDENVN-----YNNDDWYFNQIEQNIQEQIETIFDLNNFVIRSKKG 343
Qy 165 ---LYAFDNNKQNLNCKTGFKNQKFKFINHSMACDLQ-----ALDYLKRYDYNIIY-SL 215
Db 344 EFTIRAFENKQNL--QIEGLKELLF-----CKQIKQKIALNVL--NGYAKNIVLKL 393
Qy 216 NSD-----BYFKLAPDIGSDPVLSSKKPKKYINDILIPDKVAQERYGKSRLENLHYKL 270

```

Db	394	RNDIKLLIYIQLKIDFEDDHIITKGGKYIN-----MYIKKKVNNNSIKHIKEILKRN	447
Qy	271	IKDLIRLPDI	281
Db	448	IESL-NTPSNV	457

Search completed: June 13, 2005, 21:09:23
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:53:46 ; Search time 161 Seconds
(without alignments)
706.258 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNALVCGNPSLKNIID.....IRLPDIKHYLKYANKR 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	294	3	AAY97215 Campyloba
2	1571	100.0	294	6	ABJ18483 Campyloba
3	823	52.4	291	3	AAY97204 Campyloba
4	823	52.4	291	6	ABJ18479 Campyloba
5	817	52.0	291	3	AAY97211 Campyloba
6	817	52.0	291	6	ABJ18481 Campyloba
7	813	51.8	291	3	AAY97210 Campyloba
8	813	51.8	291	6	ABJ18480 Campyloba
9	804	51.2	291	3	AAY97212 Campyloba
10	804	51.2	291	6	ABJ18485 Campyloba
11	804	51.2	291	6	ABJ18484 Campyloba
12	804	51.2	291	6	ABJ18482 Campyloba
13	680.5	43.3	430	2	AAY45221 Campyloba
14	673.5	42.9	303	6	ABR40195 Campyloba
15	128.5	8.2	2013	3	AAB18265 Plasmodi
16	124.5	7.9	1398	3	AAB18292 Plasmodi
17	123	7.8	761	3	AAB16529 Bacteriop
18	121.5	7.7	1182	3	AAB18288 Plasmodi
19	119	7.6	225	6	ABU25016 Protein e
20	116.5	7.4	353	2	AAY19859 B. burgdo
21	116.5	7.4	373	2	AAY19858 B. burgdo
22	116.5	7.4	1817	3	AAB18301 Plasmodi
23	116	7.4	726	5	ABB77613 AmEPV NTP
24	115	7.3	295	5	ABB09489 AmEPV fir
25	115	7.3	1121	3	AAB18241 Plasmodi

ALIGNMENTS

RESULT 1
AAY97215

ID AAY97215 standard; protein; 294 AA.

XX AAY97215;

XX 12-SEP-2003 (revised)

DT 22-DEC-2000 (first entry)

XX

DE Campylobacter jejuni NCTC11168 alpha-2,3-sialyltransferase.

XX

KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;

KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;

KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;

KW Sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;

KW immunity; immunogen; ganglioside.

XX

OS Campylobacter jejuni; NCTC11168.

XX

PN WO200046379-A1.

XX

PD 10-AUG-2000.

XX

PF 01-FEB-2000; 2000WO-CA000086.

XX

PR 01-FEB-1999; 99US-0118213P.

PR

31-JAN-2000; 2000US-00495406.

XX

PA (CANA) NAT RES COUNCIL CANADA.

XX

PI Gilbert M, Wakarchuk WW;

XX

DR WPI; 2000-524418/47.

XX

PT Novel glycosyltransferase polypeptides and polynucleotides useful for

PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic

PT reagents and as immunogen for producing antibodies.

XX

PS Disclosure; Page 97-98; 120pp; English.

XX

CC A reaction mixture for the synthesis of a sialylated oligosaccharide is

CC useful for synthesizing sialylated oligosaccharides such as ganglioside,

CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and

CC nucleic acids that encode them are useful for studies of the pathogenesis

CC mechanisms of organisms that synthesize ganglioside mimics, such as C.

CC jejuni and the nucleic acids are used as probes to study expression of

CC

CC

genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analyzing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantity, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems. (Updated on 12-SEP-2003 to standardise OS field)

XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 1571; DB 3; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-135;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVKYVFNPFVFFE 60
DB 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVKYVFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKFFA 120

QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180

QY 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

QY 241 INDILIPDKYAQERYYGKSKLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
DB 241 INDILIPDKYAQERYYGKSKLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 2
ABJ18483
ID ABJ18483 standard; protein; 294 AA.
XX
AC ABJ18483;
XX
DT 07-FEB-2003 (first entry)
XX
DE Campylobacter jejuni bifunctional sialtransferase catII #7.
XX
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.

XX
OS Campylobacter jejuni.
XX
XX WO200274942-A2.
XX
XX 26-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-CA000229.
XX
XX 21-MAR-2001; 2001US-00816028.
XX
XX (CANADA) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;

XX
DR WPI; 2003-040554/03.
XX
PT New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Claim 5; Page 98; 107pp; English.
XX
CC The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an acyltransferase; glycosyltransferase; GalNAc (N-Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present amino acid sequence represents a Campylobacter jejuni protein of the invention
XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 1571; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-135;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVKYVFNPFVFFE 60
DB 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVKYVFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKFFA 120

QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180

QY 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

QY 241 INDILIPDKYAQERYYGKSKLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
DB 241 INDILIPDKYAQERYYGKSKLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 3
AA97204
ID AA97204 standard; protein; 291 AA.
XX
AC AA97204;
XX
DT 12-SEP-2003 (revised)
DT 22-DEC-2000 (first entry)
XX
XX Campylobacter jejuni OH4384 CstII sialyltransferase.
XX
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside.
XX
OS Campylobacter jejuni; OH4384.
XX
XX WO200046379-A1.
XX
XX 10-AUG-2000.

XX PF 01-FEB-2000; 2000WO-CA000086.
XX PR 01-FEB-1999; 99US-0118213P.
XX PR 31-JAN-2000; 2000US-00495406.
XX PA (CANA) NAT RES COUNCIL CANADA.
XX PI Gilbert M, Wakarchuk WW;
XX WPI; 2000-524418/47.
XX DR N-PSDB; AAA53721, AAA53720.
XX PT Novel glycosyltransferase polypeptides and polynucleotides useful for
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
PT reagents and as immunogen for producing antibodies.
XX PS Claim 13; Page 91-92; 120pp; English.
XX CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
CC useful for synthesising sialylated oligosaccharide such as ganglioside,
CC lyso ganglioside or their mimics. Glycosyltransferases are useful for
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analyzing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyse the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity, compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
CC (Updated on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 291 AA;
Query Match 52.4%; Score 823; DB 3; Length 291;
Best Local Similarity 54.9%; Pred. No. 4.6e-67;
Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;
QY 8 LVCNGSPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFGKDVKYVFFNPFVFFQYYSKK 67
DB 5 ILAGNGPSLKEIDYSLRPNDVDFVRCNQFYFEDRYFGKDVKYVFFNPFVFFQYYSKK 64
QY 68 LIQNEEYNIENIVCSTINLEYIDGQFQVDFNPFYFSDAFLGHEIIKKLKDFFAYIKYNEI 127
DB 65 LIQNEETELIMCSNTYNAHLENENFVKTFYDFPDHGLGYDFKQKDFNAYFKPHEI 124
QY 128 YNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLKNCTGFKNQKF 187
DB 125 YFNQRITSGVYMCATAVALGYKSIYISGIDFYQNGSS-YAFDTQKKNLLKLPNFKNDNS 183
QY 188 KFINHSMACDLOALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
DB 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIQEK-NNTYKD 242
QY 244 ILIPDKYAQERYG-----KKSRLKENLHYKIKDLRLPSDIKHYLKEK 288
DB 243 ILIPSSSEA----YGFKSNINFRKIKIKENIYYKLIKDLRLPSDIKHYPKGK 291
RESULT 4
ID ABJ18479 standard; protein; 291 AA.

XX ABJ18479;
XX DT 07-FEB-2003 (first entry)
XX DE Campylobacter jejuni bifunctional sialtransferase catII #1.
XX KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX OS Campylobacter jejuni.
XX PN WO200274942-A2.
XX PD 26-SEP-2002.
XX PF 22-FEB-2002; 2002WO-CA000229.
XX PR 21-MAR-2001; 2001US-00816028.
XX PA (CANA) NAT RES COUNCIL CANADA.
XX PI Gilbert M, Wakarchuk WW;
XX WPI; 2003-040554/03.
XX N-PSDB; ABT13666.
XX PT New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX PS Claim 5; Page 96; 107pp; English.
XX CC The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present amino acid sequence represents a Campylobacter
CC jejuni protein of the invention
XX SQ Sequence 291 AA;
Query Match 52.4%; Score 823; DB 6; Length 291;
Best Local Similarity 54.9%; Pred. No. 4.6e-67;
Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;
QY 8 LVCNGSPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFGKDVKYVFFNPFVFFQYYSKK 67
DB 5 ILAGNGPSLKEIDYSLRPNDVDFVRCNQFYFEDRYFGKDVKYVFFNPFVFFQYYSKK 64
QY 68 LIQNEEYNIENIVCSTINLEYIDGQFQVDFNPFYFSDAFLGHEIIKKLKDFFAYIKYNEI 127
DB 65 LIQNEETELIMCSNTYNAHLENENFVKTFYDFPDHGLGYDFKQKDFNAYFKPHEI 124
QY 128 YNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLKNCTGFKNQKF 187
DB 125 YFNQRITSGVYMCATAVALGYKSIYISGIDFYQNGSS-YAFDTQKKNLLKLPNFKNDNS 183
QY 188 KFINHSMACDLOALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
DB 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIQEK-NNTYKD 242

CC Acetyl(galactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present amino acid sequence represents a Campylobacter
 CC jejuni protein of the invention
 XX Sequence 291 AA;
 SQ

Query Match 52.0%; Score 817; DB 6; Length 291;
 Best Local Similarity 54.3%; Pred. No. 1.6e-66;
 Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;
 QY 8 LVCNGPSLKNIDYKRLPKQDFVFCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 67
 DB 5 ILAGNGPSLKEIDYSLRPNDVFCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 64
 QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELFSDAFLGHEIIKKLDFPFIYKNEI 127
 DB 65 LIQNEEYETELIMCSNFQAHLENQNFVKTDFYDFDAHLGYDFKQKEFNAYFKFHEI 124
 QY 128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFRNQKF 187
 DB 125 YFNQRTITSGVYMCATVALGYKSIYISGIDFYQNGSS-YAFDTKQKLLKLPNFKDINS 183
 QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDPVLSPKPKKYIND 243
 DB 184 HYIGHSKNTDIALKLEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAQERYYG-----KKSRLEKENLHYKLIKDLRLPSDIKHLYKEK 288
 DB 243 ILIPSSSEA----YGFKSNINFKKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 7
 AAY97210
 ID AAY97210 standard; protein; 291 AA.
 XX
 AC AAY97210;
 XX
 DT 12-SEP-2003 (revised)
 DT 22-DEC-2000 (first entry)
 XX
 DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
 XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside.
 XX
 OS Campylobacter jejuni; O:10 serotype.
 XX
 XX WO200046379-A1.
 XX
 PD 10-AUG-2000.
 XX
 XX 01-FEB-2000; 2000WO-CA000086.
 XX
 XX 01-FEB-1999; 99US-0118213P.
 PR 31-JAN-2000; 2000US-00495406.
 XX
 XX (CANA) NAT RES COUNCIL CANADA.
 PA
 XX Gilbert M, Wakarchuk WW;
 PI
 XX WPI; 2000-524418/47.
 DR
 DR N-PSDB; AAA53724.

XX Novel glycosyltransferase polypeptides and polynucleotides useful for
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
 PT reagents and as immunogen for producing antibodies.
 XX
 PS Claim 13; Page 92-94; 120pp; English.
 XX
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
 CC useful for synthesising sialylated oligosaccharides such as ganglioside,
 CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesise ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analysing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing
 CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyse the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems.
 CC (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 291 AA;
 Query Match 51.8%; Score 813; DB 3; Length 291;
 Best Local Similarity 54.3%; Pred. No. 3.8e-66;
 Matches 159; Conservative 46; Mismatches 70; Indels 18; Gaps 6;
 QY 8 LVCNGPSLKNIDYKRLPKQDFVFCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 67
 DB 5 ILAGNGPSLKEIDYSLRPNDVFCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 64
 QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELFSDAFLGHEIIKKLDFPFIYKNEI 127
 DB 65 LIQNEEYETELIMCSNFQAHLENQNFVKTDFYDFDAHLGYDFKQKEFNAYFKFHEI 124
 QY 128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKTGFRNQKF 187
 DB 125 YLNQRTITSGVYMCATVALGYKSIYISGIDFYQNGSS-YAFDTKQENLLKLPDFKNDRS 183
 QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDPVLSPKPKKYIND 243
 DB 184 HYIGHSKNTDIALKLEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAQERYYG-----KKSRLEKENLHYKLIKDLRLPSDIKHLYKEK 288
 DB 243 ILIPSSSEA----YGFKSNINFKKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 8
 ABJ18480
 ID ABJ18480 standard; protein; 291 AA.
 XX
 AC ABJ18480;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Campylobacter jejuni bifunctional sialtransferase cstII #2.
 XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAC transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

KW	ganglioside mimetics; inflammation; tumour metastasis.	DT	12-SEP-2003 (revised)	
XX		DT	22-DEC-2000 (first entry)	
OS		XX		
XX	Campylobacter jejuni.	DE		
XX	WO200274942-A2.	XX	Campylobacter jejuni O:19 serotype CstII sialyltransferase.	
XX		XX		
PD	26-SEP-2002.	KW	Biosynthetic locus; biosynthesis; lipid A biosynthesis; acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;	
XX		KW	Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;	
XX	22-FEB-2002; 2002WO-CA000229.	KW	sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody; immunity; immunogen; ganglioside.	
XX		XX		
XX	21-MAR-2001; 2001US-00816028.	OS	Campylobacter jejuni; O:19 serotype.	
XX	(CANA) NAT RES COUNCIL CANADA.	XX		
XX	Gilbert M, Wakarchuk WW;	XX	WO200046379-A1.	
XX		XX		
DR	WPI; 2003-040554/03.	PD	10-AUG-2000.	
DR	N-PSDB; ABT13667.	XX		
XX		XX	01-FEB-2000; 2000WO-CA000086.	
XX		XX	01-FEB-1999; 99US-0118213P.	
PT	New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.	PR	31-JAN-2000; 2000US-00495406.	
PT		XX	(CANA) NAT RES COUNCIL CANADA.	
XX		PA		
PS	Claim 5; Page 97; 107pp; English.	XX		
XX		PI	Gilbert M, Wakarchuk WW;	
CC	The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acyltransferase; glycosyltransferase; GalNAC (N-Acetyl-galactosamine) transferase; galactosyltransferase;	XX	WPI; 2000-524418/47.	
CC	sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present amino acid sequence represents a Campylobacter jejuni protein of the invention	DR	N-PSDB; AA53726.	
XX		XX		
SQ	Sequence 291 AA;	PT	Novel glycosyltransferase polypeptides and polynucleotides useful for biosynthesis of ganglioside and ganglioside mimics, as diagnostic reagents and as immunogen for producing antibodies.	
		XX	Disclosure; Page 96-97; 120pp; English.	
		CC	A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesising sialylated oligosaccharide such as ganglioside, lyso-ganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and nucleic acids that encode them are useful for studies of the pathogenesis mechanisms of organisms that synthesize ganglioside mimics, such as C. jejuni and the nucleic acids are used as probes to study expression of genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analyzing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantities, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems. (Updated on 12-SEP-2003 to standardise OS field)	
		XX		
		SQ	Sequence 291 AA;	
			Query Match	51.2%; Score 804; DB 3; Length 291;
			Best Local Similarity	53.9%; Pred. No. 2.6e-65;
			Matches	158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;
QY	8 LVCGGPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFGKDVYVFNPFVFFQYTSKK 67	QY	8 LVCGGPSLKNIIDYKRLPKQDFVRCNQFYFEDRYFGKDVYVFNPFVFFQYTSKK 67	
Db	5 IIAAGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKFKAFFVYVNPGLFFQYTLKH 64	Db	5 IIAAGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKFKAFFVYVNPGLFFQYTLKH 64	
QY	68 LIQNEEYNIENIVCSTINLEYIDGQFVDNPFELVFSDAFLGHEIIKKLDFFAYIKYNEI 127	QY	68 LIQNEEYNIENIVCSTINLEYIDGQFVDNPFELVFSDAFLGHEIIKKLDFFAYIKYNEI 127	
Db	65 LIQNEEYETELIMCSNYNOAHLENENFVKTFYDFPDHGLGYOFFKQKGFNAYFKFHEI 124	Db	65 LIQNEEYETELIMCSNYNOAHLENENFVKTFYDFPDHGLGYOFFKQKGFNAYFKFHEI 124	
QY	128 YNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKNLLNKCTGFRKNQF 187	QY	128 YNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKNLLNKCTGFRKNQF 187	
Db	125 YLNQRITSGVYMCATAVALGYKEILSGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183	Db	125 YLNQRITSGVYMCATAVALGYKEILSGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183	
QY	188 KFINHSMACDLQALDYLKDYDNVNIYSL--NS--DEYFKLAPDIGSDFVLSPKPKYIND 243	QY	188 KFINHSMACDLQALDYLKDYDNVNIYSL--NS--DEYFKLAPDIGSDFVLSPKPKYIND 243	
Db	184 HYIGHSKNTDIKALEFLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIIQEK--NNYTKD 242	Db	184 HYIGHSKNTDIKALEFLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIIQEK--NNYTKD 242	
QY	244 ILIPDKVAQERYYG-----KKSRLKENLHYKLIKDLRLSPDINKHYLKEK 288	QY	244 ILIPDKVAQERYYG-----KKSRLKENLHYKLIKDLRLSPDINKHYLKEK 288	
Db	243 ILIPSESA----YKFSKKNINFKKIKENIYYKLIKDLRLSPDINKHYFKGK 291	Db	243 ILIPSESA----YKFSKKNINFKKIKENIYYKLIKDLRLSPDINKHYFKGK 291	
RESULT 9				
AAAY97212				
ID	AAAY97212 standard; protein; 291 AA.			
XX				
AC	AAAY97212;			
XX				

QY 128 YNRQRITSGVVMCATAVAGYKSIYISGIDFYQDTNNLYAFDNNKNLKNKCTGFKNQKF 187
 Db 125 YFNQRITSGVVMCAVALGKKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183
 QY 188 KFINHSMACDQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGDFVLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEFLEKTKIKLYCLCPNSLLANFIELAPLNSNFIQEK-NNYTKD 242
 QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLIKDLRLPSDIKHYYLKEK 288
 Db 243 ILIPSSSEA----YGFKSKNINFKIKIKENYVYKLIKDLRLPSDIKHYPKKG 291

RESULT 10
 ABJ18485
 ID ABJ18485 standard; protein; 291 AA.
 XX
 AC ABJ18485;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Campylobacter jejuni bifunctional sialtransferase catII #6.
 XX
 KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 FN WO200274942-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-CA000229.
 XX
 PR 21-MAR-2001; 2001US-00816028.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2003-040554/03.
 DR N-PSDB; ABT13671.
 XX
 PT New glycosyltransferases from Campylobacter, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.
 XX
 PS Disclosure; Page 99; 107pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetylgalactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present amino acid sequence represents a Campylobacter
 CC jejuni protein of the invention

SQ Sequence 291 AA;

Query Match 51.2%; Score 804; DB 6; Length 291;

Best Local Similarity 53.9%; Pred. No. 2.6e-65; Indels 18; Gaps 6;
 Matches 158; Conservative 45; Mismatches 72;

QY 8 LVCNGPSPSLKNIDYKPLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFPQYYSKK 67
 Db 5 IIAAGPSPSLKEIDYSRLPNDVFRCNQFYFEDRYFVGKDVKYVFNPFVFPQYYSKK 64
 QY 68 LIQNEBYNIENIVCSTINLEYIDGFOFVDNFELYSFSDAFGLGHEIIKKLQDFRAYIKYNEI 127
 Db 65 LIQNEBYTELIMCSNYNQAHLENENFVKTFYDFDAHLGIDYFQOLKEFNAYFKFHEI 124
 QY 128 YNRQRITSGVVMCATAVAGYKSIYISGIDFYQDTNNLYAFDNNKNLKNKCTGFKNQKF 187
 Db 125 YFNQRITSGVVMCAVALGKKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183
 QY 188 KFINHSMACDQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGDFVLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEFLEKTKIKLYCLCPNSLLANFIELAPLNSNFIQEK-NNYTKD 242
 QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLIKDLRLPSDIKHYYLKEK 288
 Db 243 ILIPSSSEA----YGFKSKNINFKIKIKENYVYKLIKDLRLPSDIKHYPKKG 291

RESULT 11

ABJ18484

ID ABJ18484 standard; protein; 291 AA.

XX AC ABJ18484;

XX DT 07-FEB-2003 (first entry)

XX DE Campylobacter jejuni bifunctional sialtransferase catII #5.

XX KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.
 XX FN WO200274942-A2.

XX PD 26-SEP-2002.
 XX PF 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.
 XX PA (CANADA) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;
 XX DR WPI; 2003-040554/03.

XX DR N-PSDB; ABT13670.
 XX PT New glycosyltransferases from Campylobacter, useful for synthesizing
 XX PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 XX PT mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Disclosure; Page 99; 107pp; English.
 XX CC The invention comprises the amino acid and coding sequences of
 XX CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 XX CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 XX CC Acetylgalactosamine) transferase; galactosyltransferase;
 XX CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 XX CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 XX CC sequences of the invention are useful for ganglioside synthesis, studying
 XX CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 XX CC expression of Campylobacter enzymes involved in the biosynthesis of
 XX CC ganglioside mimetics that can mask the pathogen's from the host's immune
 XX CC system. The C. jejuni oligosaccharides of the invention may be used as

CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present amino acid sequence represents a Campylobacter
CC jejuni protein of the invention
XX
SQ Sequence 291 AA;

Query Match 51.2%; Score 804; DB 6; Length 291;
Best Local Similarity 53.9%; Pred. No. 2.6e-65;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCGGSLKNIIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFFQYYSKK 67
Db 5 ILAGGSLKEIDYSLRPNDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFFQYYSKK 64

QY 68 LIQNEEYNINIVCSTINLEYIDGQFVFNELFYSDFALGHEIIKKLDPFFAYIKYNEI 127
Db 65 LIQNEEYETELIMCSYNAHLENENFVKTFYDFPDHGLGYDFKQLKEFNAYFKFHEI 124

QY 128 YNQRITSGVYMCATAVALGYSIYISGIDFYDTNNLYAFDNNKGNLLKCTGFKNQKF 187
Db 125 YFNQRITSGVYMCATAVALGYSIYISGIDFYDTNNLYAFDNNKGNLLKCTGFKNQKF 183

QY 188 KFINHSMACDLQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGDFVLSKPKKYIND 243
Db 184 HYGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242

QY 244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDIKHLYKEK 288
Db 243 ILIPSSSEA----YGFKSKNINFKIKENYVYKLIKDLRLPSDIKHLYFGKG 291

RESULT 12
ABJ18482
ID ABJ18482 standard; protein; 291 AA.
XX
AC ABJ18482;
XX
DT 07-FEB-2003 (first entry)
XX
DE Campylobacter jejuni bifunctional sialtransferase cstII #4.
XX
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
PN WO200274942-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA000229.
XX
PR 21-MAR-2001; 2001US-00816028.
XX
PA (CANADA) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Wakarchuk WW;
XX
XX WPI; 2003-040554/03.
XX
XX N-PSDB; ABT13669.
XX
XX New glycosyltransferases from Campylobacter, useful for synthesizing
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis
XX PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
XX PS Disclosure; Page 99; 107pp; English.
XX
XX The invention comprises the amino acid and coding sequences of
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention

CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present amino acid sequence represents a Campylobacter
CC jejuni protein of the invention
XX
SQ Sequence 291 AA;

Query Match 51.2%; Score 804; DB 6; Length 291;
Best Local Similarity 53.9%; Pred. No. 2.6e-65;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCGGSLKNIIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFFQYYSKK 67
Db 5 ILAGGSLKEIDYSLRPNDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFFQYYSKK 64

QY 68 LIQNEEYNINIVCSTINLEYIDGQFVFNELFYSDFALGHEIIKKLDPFFAYIKYNEI 127
Db 65 LIQNEEYETELIMCSYNAHLENENFVKTFYDFPDHGLGYDFKQLKEFNAYFKFHEI 124

QY 128 YNQRITSGVYMCATAVALGYSIYISGIDFYDTNNLYAFDNNKGNLLKCTGFKNQKF 187
Db 125 YFNQRITSGVYMCATAVALGYSIYISGIDFYDTNNLYAFDNNKGNLLKCTGFKNQKF 183

QY 188 KFINHSMACDLQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGDFVLSKPKKYIND 243
Db 184 HYGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242

QY 244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDIKHLYKEK 288
Db 243 ILIPSSSEA----YGFKSKNINFKIKENYVYKLIKDLRLPSDIKHLYFGKG 291

RESULT 13
AA45221
ID AA45221 standard; protein; 430 AA.
XX
AC AA45221;
XX
DT 05-JAN-2000 (first entry)
XX
DE Campylobacter jejuni alpha-2,3-sialyltransferase protein.
XX
KW Campylobacter jejuni; alpha-2,3-sialyltransferase; cst-I; acceptor;
KW lipopolysaccharide; galactose residue; sialic acid molecule.
XX
OS Campylobacter jejuni.
XX
PN WO9949051-A1.
XX
PD 30-SEP-1999.
XX
PF 22-MAR-1999; 99WO-CA000238.
XX
PR 20-MAR-1998; 98US-0078891P.
XX
PR 18-MAR-1999; 99US-00272960.
XX
PA (CANADA) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Wakarchuk WW;
XX
XX WPI; 1999-601216/51.
XX
XX N-PSDB; AA225693.
XX
XX New sialyltransferases useful for adding sialyl residues to acceptor
XX PT molecules.

```

PS Claim 27; Fig 2; 47pp; English.
XX
CC The present sequence represents Campylobacter jejuni alpha-2,3-
CC sialyltransferase which is encoded by the cst-I gene. The alpha-2,3-
CC sialyltransferase protein is useful for producing desired carbohydrate
CC structures by contacting the acceptor molecule (which has a terminal
CC galactose residue) with an activated sialic acid molecule. The terminal
CC galactose residue is linked to a second residue (Glc or a GlcNac, or
CC GlcNac or Galnac) in the acceptor molecule through a beta-1,3 or beta-1,4
CC linkage, respectively. The activated sialic acid is CMP-Neu5Ac. The
CC polynucleotides and polypeptides facilitate the improved production of
CC desired structures and nucleic acids encoding sialyltransferases
XX
SQ Sequence 430 AA;
    Query Match      43.3%; Score 680.5; DB 2; Length 430;
    Best Local Similarity 49.7%; Pred. No. 9.1e-54;
    Matches 145; Conservative 49; Mismatches 75; Indels 23; Gaps 8;
QY 1 MSMMI-NALVCGNGPSLKNIDYKRLPKQFVFCRCNQFYFEDRYFVGKDVYFNPFPVFF 59
DB 12 VSKMQNIIAGNGPSLKNINIKLPREYDVFRCNQFYFEDRYFVGKDVYFNPFPVFF 71
QY 60 EGYTSKLLIQNEEYNIENIVCSTINLEYIDGFOVDNEFLYSDAFLGHEIILKLDKFF 119
DB 72 QVHTAKQLIKNIEYEIKNIFCSFTNPFIESNDLFHQFYNFPFDKLGVEIENLKFEY 131
QY 120 AYIKNEYINRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNKC 179
DB 132 AYIKNEYINRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNKC 190
QY 180 TGFKNQKPKTN-HSMACDQALDYLKMKRYDVNIYSLNSD-----EYFKLAPDIGSDPVL 234
DB 191 PGIKD--FKPSNCHSKSEYDEALKLSIYKVNIALCDSILANHPPLSININNNFTLE 248
QY 235 KPKKYINDILIPDKYAQERYGKSRLEN-----LHYK--LIKDL 274
DB 249 NKHNSINDILLTNTPGVSFY--KNQLKADNKIMLNFYILHSLKONLIKPL 298

RESULT 14
ABR40195
ID ABR40195 standard; protein; 303 AA.
AC ABR40195;
XX
XX 23-JUL-2003 (first entry)
DE alpha-2,3/alpha-2,8-sialyltransferase #3.
XX alpha-2,3/alpha-2,8-sialyltransferase; enzyme; sialic acid.
XX Pasteurella multocida.
XX WO2003027297-A1.
XX 03-APR-2003.
XX 26-SEP-2002; 2002WO-JP009907.
XX 26-SEP-2001; 2001JP-00292796.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX Endo T, Koizumi S;
XX WPI; 2003-393339/37.
XX N-PSDB; ACC71694.
XX Alpha 2,3/alpha 2,8 sialyltransferase from Pasteurella multocida for
XX production of sialic acid-containing complex sugars.
XX
PS Claim 9; Page 60-62; 66pp; Japanese.
XX
CC The present invention relates to a method for producing alpha-2,3/alpha-
CC 2,8-sialyltransferase. alpha-2,3/alpha-2,8-sialyltransferase can be used
CC in a method for producing sialic acid-containing complex sugars from an
CC oligosaccharide substrate using a culture of a microorganism transformed
CC with DNA encoding alpha-2,3/alpha-2,8-sialyltransferase, or an extract of
CC the culture. The present sequence is a protein sequence for alpha-
CC 2,3/alpha-2,8-sialyltransferase from Pasteurella multocida
XX
SQ Sequence 303 AA;
    Query Match      42.9%; Score 673.5; DB 6; Length 303;
    Best Local Similarity 46.8%; Pred. No. 2.6e-53;
    Matches 137; Conservative 47; Mismatches 94; Indels 15; Gaps 7;
QY 8 LVCGNGPSLKNIDYKRLPKQFVFCRCNQFYFEDRYFVGKDVYFNPFPVFFEQYTSKK 67
DB 14 IVAGNESLSQIDYRLLPKNYDVFRCNQFYFEDRYFVGKDVYFNPFPVFFEQYTSKK 73
QY 68 LIQNEEYNIENIVCSTINLEYIDGFOVDNEFLYSDAFLGHEIILKLDKFPAYIKNE 126
DB 74 LKRNNEYFVDNVILSSFNHPTVD-LEKSQKIQALFIDVINGYEKLSKLTAFDVLRYKE 132
QY 127 IYNRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNKC 186
DB 133 LYENRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNKC 192
QY 187 FKFINHSMACDQALDYLKMKRYDVNIYSLNS-----DHYFKLAPDIGSD----FVLSKKPK 238
DB 193 TLFYSYHKKDIDLEALSFLQOHYHVNYSISPMSPSKHFPF-PTVEDDCETTFVAPLK-E 250
QY 239 KYINDILIPDKYAQER---YYGKSRLENLHYKLIKDLIRLPSDIKHYLKEK 288
DB 251 NYINDILLPPHFVYKLGITVSKSRFHSNLIVRLIRDLKLSALKHYLKEK 303

RESULT 15
AAB18265
ID AAB18265 standard; protein; 2013 AA.
XX
XX AAB18265;
XX
XX 07-NOV-2000 (first entry)
DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:122.
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoacide; infection; insecticide.
XX Plasmodium falciparum.
XX WO200025728-A2.
XX 11-MAY-2000.
XX 05-NOV-1999; 99WO-US026796.
XX 05-NOV-1998; 98US-0107131P.
XX (HOFF/) HOFFMAN S.
XX (CARU/) CARUCCI D.
XX (GARD/) GARDNER M.
XX (VENT/) VENTER J C.
XX Hoffman S, Carucci D, Gardner M, Venter JC;
XX WPI; 2000-365347/31.
XX
XX Proteins encoded by chromosome 2 of the human malarial parasite,
XX Plasmodium falciparum, useful as antimalarial vaccines and in the
XX diagnosis of P.falciparum infection.
XX

```

Disclosure; Page 285-291; 577pp; English.

The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against *P. falciparum* infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with *P. falciparum*. Furthermore, (I) (especially when they are inflix or secreted or membrane proteins) can aid the identification of drugs to treat or prevent *P. falciparum* infection, or they can be used to identify drug resistance in *P. falciparum*. Sequencing of the *Plasmodium* chromosome 2 and the subsequent identifying of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAT0078 to AAT70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

Sequence 2013 AA;

Search completed: June 13, 2005, 21:06:23
Job time : 164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 21:00:06 ; Search time 43 Seconds
(without alignments)
510.391 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNINALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKR 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgm2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgm2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgm2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgm2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgm2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1571	100.0	294	4	US-09-495-406-10
2	1571	100.0	294	4	US-09-816-028A-10
3	1571	100.0	294	4	US-10-303-162-10
4	1571	100.0	294	4	US-10-303-134-10
5	823	52.4	291	4	US-09-495-406-3
6	823	52.4	291	4	US-09-816-028A-3
7	823	52.4	291	4	US-10-303-162-3
8	823	52.4	291	4	US-10-303-134-3
9	817	52.0	291	4	US-09-495-406-7
10	817	52.0	291	4	US-09-816-028A-7
11	817	52.0	291	4	US-10-303-162-7
12	817	52.0	291	4	US-10-303-134-7
13	813	51.8	291	4	US-09-495-406-5
14	813	51.8	291	4	US-09-816-028A-5
15	813	51.8	291	4	US-10-303-162-5
16	813	51.8	291	4	US-10-303-134-5
17	804	51.2	291	4	US-09-495-406-9
18	804	51.2	291	4	US-09-816-028A-9
19	804	51.2	291	4	US-09-816-028A-12
20	804	51.2	291	4	US-09-816-028A-14
21	804	51.2	291	4	US-10-303-162-9
22	804	51.2	291	4	US-10-303-162-12
23	804	51.2	291	4	US-10-303-162-14
24	804	51.2	291	4	US-10-303-134-9
25	804	51.2	291	4	US-10-303-134-12
26	804	51.2	291	4	US-10-303-134-14
27	680.5	43.3	322	4	US-09-495-406-34

28	680.5	43.3	322	4	US-09-816-028A-48	Sequence 48, Appl
29	680.5	43.3	322	4	US-10-303-162-48	Sequence 48, Appl
30	680.5	43.3	322	4	US-10-303-134-48	Sequence 48, Appl
31	680.5	43.3	430	4	US-09-272-960-2	Sequence 2, Appl
32	680.5	43.3	430	4	US-10-058-636-2	Sequence 2, Appl
33	441.5	28.1	231	4	US-09-495-406-35	Sequence 35, Appl
34	441.5	28.1	231	4	US-09-272-960-5	Sequence 5, Appl
35	441.5	28.1	231	4	US-09-816-028A-49	Sequence 49, Appl
36	441.5	28.1	231	4	US-10-058-636-5	Sequence 5, Appl
37	441.5	28.1	231	4	US-10-303-162-49	Sequence 49, Appl
38	441.5	28.1	231	4	US-10-303-134-49	Sequence 49, Appl
39	112	7.1	1169	4	US-09-583-110-4409	Sequence 4409, Ap
40	112	7.1	1169	4	US-09-107-433-4847	Sequence 4847, Ap
41	104	6.6	367	4	US-09-861-451A-46	Sequence 46, Appl
42	104	6.6	898	1	US-08-465-995A-4	Sequence 4, Appl
43	104	6.6	898	2	US-08-465-994C-4	Sequence 4, Appl
44	104	6.6	898	2	US-08-966-145-4	Sequence 4, Appl
45	104	6.6	920	1	US-08-101-593-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-495-406-10
; Sequence 10, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168

Query Match	100.0%	Score 1571;	DB 4;	Length 294;	
Best Local Similarity	100.0%	Pred. No. 1e-153;			
Matches 294;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	MSNNINALVCGNGPSLKNIIDYKELPKQFVFCRNCQFVEDRYFVGKDVKYVFNPFVFFE	60		
Db	1	MSNNINALVCGNGPSLKNIIDYKELPKQFVFCRNCQFVEDRYFVGKDVKYVFNPFVFFE	60		
Qy	61	QYTSKKLIQNEBYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLKDFFA	120		
Db	61	QYTSKKLIQNEBYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLKDFFA	120		
Qy	121	YIKYNIYNQRITSGVYMCATAVALGYKSIYISGIDFYODTNNLYAFDNNKKNLNKCT	180		
Db	121	YIKYNIYNQRITSGVYMCATAVALGYKSIYISGIDFYODTNNLYAFDNNKKNLNKCT	180		
Qy	181	GFKNQKFKETINSHMACDLQALDYLMKRYDVNLYSLNSDEYFKLAPDIGSDFVLSKKPKKY	240		
Db	181	GFKNQKFKETINSHMACDLQALDYLMKRYDVNLYSLNSDEYFKLAPDIGSDFVLSKKPKKY	240		
Qy	241	INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDILRLPSDIKHYLKEKYANKR	294		
Db	241	INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDILRLPSDIKHYLKEKYANKR	294		

QY 61 QYTSKLIQNEEYNIENIVCSTINLEYIDGQFQVDNFELYSDFDAFLGHEIIKKLKDFFA 120
 Db 61 QYTSKLIQNEEYNIENIVCSTINLEYIDGQFQVDNFELYSDFDAFLGHEIIKKLKDFFA 120
 QY 121 YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 Db 121 YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 QY 181 GFKNQKFKFNHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIDGSDFLSKPKKY 240
 Db 181 GFKNQKFKFNHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIDGSDFLSKPKKY 240
 QY 241 INDILIPDKYAQERYGKSKRLKLNHLYKLIKOLIRLPSDIKHYLKEKYANKR 294
 Db 241 INDILIPDKYAQERYGKSKRLKLNHLYKLIKOLIRLPSDIKHYLKEKYANKR 294

RESULT 5

US-09-495-406-3
 ; Sequence 3, Application US/09495406
 ; Patent No. 6503744
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; FILE REFERENCE: Gangliosides and Ganglioside Mimics
 ; CURRENT APPLICATION NUMBER: US/09/495,406
 ; CURRENT FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Campylobacter jejuni
 ; US-09-495-406-3

Query Match 52.4%; Score 823; DB 4; Length 291;
 Best Local Similarity 54.9%; Pred. No. 1.4e-76;
 Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 67
 Db 5 IIAGNGPSLKEIDYSLRPNDVFRQCQFYFEDKYLYGKKCAVFNYPILFFQYVTLKH 64
 QY 68 LIQNEEYNIENIVCSTINLEYIDGQFQVDNFELYSDFDAFLGHEIIKKLKDFFAYIKNEI 127
 Db 65 LIQNEEYETELIMCSNYNOAHLENENFVKTFDYDPAHLGDFPKQLKDFNAYPKFHEI 124
 QY 128 YNRQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCTGFKNQKF 187
 Db 125 YFNQRITSGVYMCATAVALGYKEIYLSGIDFYQNGSS-YAFDTKQKLLKLANFNKDNS 183
 QY 188 KFINHSMACDLQALDYLKRYDNIYSL--NS--DEYFKLAPDIDGSDFLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEPLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAQERYG-----KKSRLKENLHLYKLIKOLIRLPSDIKHYLKEK 288
 Db 243 ILIPSSA----YGFKSKNINFKKIKENIYKLIKOLIRLPSDIKHYFKGK 291

RESULT 6

US-09-816-028A-3
 ; Sequence 3, Application US/09816028A
 ; Patent No. 6699705
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; FILE REFERENCE: Gangliosides and Ganglioside Mimics
 ; CURRENT APPLICATION NUMBER: US/09/816,028A
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (CstII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-09-816-028A-3

Query Match 52.4%; Score 823; DB 4; Length 291;
 Best Local Similarity 54.9%; Pred. No. 1.4e-76;
 Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;
 QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 67
 Db 5 IIAGNGPSLKEIDYSLRPNDVFRQCQFYFEDKYLYGKKCAVFNYPILFFQYVTLKH 64
 QY 68 LIQNEEYNIENIVCSTINLEYIDGQFQVDNFELYSDFDAFLGHEIIKKLKDFFAYIKNEI 127
 Db 65 LIQNEEYETELIMCSNYNOAHLENENFVKTFDYDPAHLGDFPKQLKDFNAYPKFHEI 124
 QY 128 YNRQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCTGFKNQKF 187
 Db 125 YFNQRITSGVYMCATAVALGYKEIYLSGIDFYQNGSS-YAFDTKQKLLKLANFNKDNS 183
 QY 188 KFINHSMACDLQALDYLKRYDNIYSL--NS--DEYFKLAPDIDGSDFLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEPLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAQERYG-----KKSRLKENLHLYKLIKOLIRLPSDIKHYLKEK 288
 Db 243 ILIPSEA----YGFKSKNINFKKIKENIYKLIKOLIRLPSDIKHYFKGK 291

RESULT 7

US-10-303-162-3
 ; Sequence 3, Application US/10303162
 ; Patent No. 6723545
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; FILE REFERENCE: Gangliosides and Ganglioside Mimics
 ; CURRENT APPLICATION NUMBER: US/10/303,162
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Campylobacter jejuni

```
;
; FEATURE:
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (CetII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-3

Query Match      52.4%; Score 823; DB 4; Length 291;
Best Local Similarity 54.9%; Pred. No. 1.4e-76;
Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQFVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 67
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 IIAGNGPSLKEIDYSLPNDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 64
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDPFAYIKYNEI 127
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 LIQNEEYETELIMCSNYNQAHLNENFVKTFFYDPDAHLGIDYFFKQLKDNFAYFKFHEI 124
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 YNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 187
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 YFNQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 183
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK--NNYTKD 242
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLKDLIRLPSDIKHYLKEK 288
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 ILIPSEA-----YGFPSKNINFKKIKENIYKLIKDLRLPSDIKHYFGKG 291
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 YNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 187
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 YFNQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 183
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK--NNYTKD 242
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLKDLIRLPSDIKHYLKEK 288
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 ILIPSEA-----YGFPSKNINFKKIKENIYKLIKDLRLPSDIKHYFGKG 291
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-303-134-3
; Sequence 3, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (CetII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-134-3

Query Match      52.4%; Score 823; DB 4; Length 291;
Best Local Similarity 54.9%; Pred. No. 1.4e-76;
Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQFVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 67
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 IIAGNGPSLKEIDYSLPNDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 64
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDPFAYIKYNEI 127
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 LIQNEEYETELIMCSNYNQAHLNENFVKTFFYDPDAHLGIDYFFKQLKDNFAYFKFHEI 124
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 YNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 187
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 YFNQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 183
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK--NNYTKD 242
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLKDLIRLPSDIKHYLKEK 288
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 ILIPSEA-----YGFPSKNINFKKIKENIYKLIKDLRLPSDIKHYFGKG 291
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-495-406-7
; Sequence 7, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
US-09-495-406-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQFVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 67
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 IIAGNGPSLKEIDYSLPNDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFQYTSKK 64
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDPFAYIKYNEI 127
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 LIQNEEYETELIMCSNYNQAHLNENFVKTFFYDPDAHLGIDYFFKQLKDNFAYFKFHEI 124
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 YNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 187
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 YFNQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQKF 183
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK--NNYTKD 242
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLKDLIRLPSDIKHYLKEK 288
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 ILIPSEA-----YGFPSKNINFKKIKENIYKLIKDLRLPSDIKHYFGKG 291
   : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-816-028A-7
; Sequence 7, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
```

```
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-09-816-028A-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY      8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGDKVYFNFPPVFFPQQYTSKK 67
DB      5 ILAGNGPSLKEIDYSLRPNDQDFVRCNQFYFEDRYFVGDKVYFNFPPVFFPQQYTSKK 64
QY      68 LIQNEEYNINIVCSTINLEYIDGQFQVDFNPFYSDAFLGHEIIKKLDKDFPFIKYNEI 127
DB      65 LIQNEEYNINIVCSTINLEYIDGQFQVDFNPFYSDAFLGHEIIKKLDKDFPFIKYNEI 124
QY      128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFKNQKF 187
DB      125 YFNQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFKNQKF 183
QY      188 KFINHSMACDLQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
DB      184 HYIGHSKNTDIKALEFLEKTYEIKLYCLPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
QY      244 ILIPDKYAORYYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
DB      243 ILIPDSEA---YGFKTKNINFKIKENIYKLIKDLIRLPSDIKHYLKEK 291

RESULT 12
US-10-303-134-7
; Sequence 7, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-134-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY      8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGDKVYFNFPPVFFPQQYTSKK 67
DB      5 ILAGNGPSLKEIDYSLRPNDQDFVRCNQFYFEDRYFVGDKVYFNFPPVFFPQQYTSKK 64
QY      68 LIQNEEYNINIVCSTINLEYIDGQFQVDFNPFYSDAFLGHEIIKKLDKDFPFIKYNEI 127
DB      65 LIQNEEYNINIVCSTINLEYIDGQFQVDFNPFYSDAFLGHEIIKKLDKDFPFIKYNEI 124
QY      128 YNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFKNQKF 187
DB      125 YFNQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFKNQKF 183
QY      188 KFINHSMACDLQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
DB      184 HYIGHSKNTDIKALEFLEKTYEIKLYCLPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
QY      244 ILIPDKYAORYYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
DB      243 ILIPDSEA---YGFKTKNINFKIKENIYKLIKDLIRLPSDIKHYLKEK 291

RESULT 11
US-10-303-162-7
; Sequence 7, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-162-7
```


Search completed: June 13, 2005, 21:10:56
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 21:09:32 ; Search time 157 Seconds
(without alignments)
717.836 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNALVCGNGPSLKNID.....IRLPDIKHVKEKYANKNR 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1571	100.0	294	9	US-09-816-028A-10
2	1571	100.0	294	14	US-10-303-161-10
3	1571	100.0	294	14	US-10-303-118-10
4	1571	100.0	294	14	US-10-303-128-10
5	1571	100.0	294	14	US-10-303-134-10
6	1571	100.0	294	14	US-10-303-162-10
7	1571	100.0	294	16	US-10-820-536-10
8	1571	100.0	294	16	US-10-845-408-10
9	1571	100.0	294	16	US-10-845-412-10
10	1571	100.0	294	16	US-10-846-219-10
11	1571	100.0	294	16	US-10-821-604-10

12	1571	100.0	294	16	US-10-847-983-10	Sequence 10, Appl
13	1571	100.0	294	16	US-10-821-573-10	Sequence 10, Appl
14	1571	100.0	294	16	US-10-850-807-10	Sequence 10, Appl
15	1571	100.0	294	16	US-10-850-125-10	Sequence 10, Appl
16	1571	100.0	294	16	US-10-830-825-10	Sequence 10, Appl
17	1571	100.0	294	17	US-10-962-334-10	Sequence 10, Appl
18	1571	100.0	294	17	US-10-830-997-10	Sequence 10, Appl
19	1571	100.0	294	17	US-10-962-235-10	Sequence 10, Appl
20	823	52.4	291	9	US-09-816-028A-3	Sequence 3, Appl
21	823	52.4	291	14	US-10-303-161-3	Sequence 3, Appl
22	823	52.4	291	14	US-10-303-118-3	Sequence 3, Appl
23	823	52.4	291	14	US-10-303-128-3	Sequence 3, Appl
24	823	52.4	291	14	US-10-303-134-3	Sequence 3, Appl
25	823	52.4	291	14	US-10-303-162-3	Sequence 3, Appl
26	823	52.4	291	16	US-10-820-536-3	Sequence 3, Appl
27	823	52.4	291	16	US-10-845-408-3	Sequence 3, Appl
28	823	52.4	291	16	US-10-845-412-3	Sequence 3, Appl
29	823	52.4	291	16	US-10-846-219-3	Sequence 3, Appl
30	823	52.4	291	16	US-10-821-604-3	Sequence 3, Appl
31	823	52.4	291	16	US-10-847-983-3	Sequence 3, Appl
32	823	52.4	291	16	US-10-821-573-3	Sequence 3, Appl
33	823	52.4	291	16	US-10-850-807-3	Sequence 3, Appl
34	823	52.4	291	16	US-10-850-125-3	Sequence 3, Appl
35	823	52.4	291	16	US-10-830-825-3	Sequence 3, Appl
36	823	52.4	291	17	US-10-962-334-3	Sequence 3, Appl
37	823	52.4	291	17	US-10-830-997-3	Sequence 3, Appl
38	823	52.4	291	17	US-10-962-235-3	Sequence 3, Appl
39	817	52.0	291	9	US-09-816-028A-7	Sequence 7, Appl
40	817	52.0	291	14	US-10-303-161-7	Sequence 7, Appl
41	817	52.0	291	14	US-10-303-118-7	Sequence 7, Appl
42	817	52.0	291	14	US-10-303-128-7	Sequence 7, Appl
43	817	52.0	291	14	US-10-303-134-7	Sequence 7, Appl
44	817	52.0	291	14	US-10-303-162-7	Sequence 7, Appl
45	817	52.0	291	16	US-10-820-536-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-10
; Sequence 10, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-09-816-028A-10

Query Match 100.0%; Score 1571; DB 9; Length 294;

Best Local Similarity 100.0%; Pred. No. 1.6e-125;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNNALVCGNGPSLKNIDYKELPKQDFRCNQYFEDRYFVGKDKVFFNPFVFF 60
|||||


```
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-128-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQNFYFEDRYFVGKDVKYVFNFPVFPE 60
Db 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQNFYFEDRYFVGKDVKYVFNFPVFPE 60
Qy 61 QYVTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKGNLNKCT 120
Db 61 QYVTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKGNLNKCT 120
Qy 121 YIKYNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKGNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKGNLNKCT 180
Qy 181 GFKNQKPKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db 181 GFKNQKPKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 6
US-10-303-162-10
; Sequence 10, Application US/10303162
; Publication No. US20030157658A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-162-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQNFYFEDRYFVGKDVKYVFNFPVFPE 60
Db 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQNFYFEDRYFVGKDVKYVFNFPVFPE 60
Qy 61 QYVTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKGNLNKCT 120
Db 61 QYVTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKGNLNKCT 120
Qy 121 YIKYNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKGNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKGNLNKCT 180
Qy 181 GFKNQKPKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db 181 GFKNQKPKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 5
US-10-303-134-10
; Sequence 10, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-134-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQNFYFEDRYFVGKDVKYVFNFPVFPE 60
```

RESULT 7
US-10-820-536-10
; Sequence 10, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-820-536-10
Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSNNALVCGNGPSLKNIDYKELPKQFDVFRNCNQYFEDRYFVGKDVYFNFVFFPE 60
Db 1 MSNNALVCGNGPSLKNIDYKELPKQFDVFRNCNQYFEDRYFVGKDVYFNFVFFPE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLKDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLKDFFA 120
Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYEYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYEYANKNR 294
RESULT 8
US-10-845-408-10
; Sequence 10, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-845-408-10
Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSNNALVCGNGPSLKNIDYKELPKQFDVFRNCNQYFEDRYFVGKDVYFNFVFFPE 60
Db 1 MSNNALVCGNGPSLKNIDYKELPKQFDVFRNCNQYFEDRYFVGKDVYFNFVFFPE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLKDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLKDFFA 120
Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYEYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYEYANKNR 294
RESULT 9
US-10-845-412-10
; Sequence 10, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-845-412-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60
DB 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120

QY 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
DB 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240

QY 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
DB 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 10
US-10-846-219-10
; Sequence 10, Application US/10846219
; Publication No. US20040219638A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-846-219-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60
DB 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120

QY 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
DB 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240

Db 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
QY 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
Db 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 11
US-10-821-604-10
; Sequence 10, Application US/10821604
; Publication No. US20040229263A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,604
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-821-604-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60
DB 1 MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYVFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLKDFFA 120

QY 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
DB 181 GFKNQKFKFINHSMACDLQALDYLMKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240

QY 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
DB 241 INDILIPDKYAQRYRYGKKSRLKENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 12
US-10-847-983-10
; Sequence 10, Application US/10847983
; Publication No. US20040229272A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-847-983-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

RESULT 13
US-10-821-573-10
; Sequence 10, Application US/10821573
; Publication No. US2004029313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: Gangliosides and Ganglioside Mimics
US-10-821-573-10

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-807-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

RESULT 14
US-10-850-807-10
; Sequence 10, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-807-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

```
Db      121  YIKNEIYNQRITSGVTCATAVALGYSIYISGIDFYQDTNNLYAFDNNKGLLNKCT 180
Qy      181  GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db      181  GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy      241  INDILIPDKYAQERYGKSKRLKLNHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db      241  INDILIPDKYAQERYGKSKRLKLNHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
```

RESULT 15

```
US-10-850-125-10
; Sequence 10, Application US/10850125
; Publication No. US20040259203A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-125-10
```

```
Query Match      100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MSNNALVCGNPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db      1  MSNNALVCGNPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFE 60
Qy      61  QYTSKGLIQNEEYNIENIVCSTINLEYIDGFQFVDNFELVFSDAPLGEIHKLDPPA 120
Db      61  QYTSKGLIQNEEYNIENIVCSTINLEYIDGFQFVDNFELVFSDAPLGEIHKLDPPA 120
Qy      121 YIKNEIYNQRITSGVTCATAVALGYSIYISGIDFYQDTNNLYAFDNNKGLLNKCT 180
Db      121 YIKNEIYNQRITSGVTCATAVALGYSIYISGIDFYQDTNNLYAFDNNKGLLNKCT 180
Qy      181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db      181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy      241 INDILIPDKYAQERYGKSKRLKLNHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db      241 INDILIPDKYAQERYGKSKRLKLNHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
```

Search completed: June 13, 2005, 21:22:52
Job time : 158 secs

This Page Blank (uspto)

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:17542800
A;Accession: B64149
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-231 <TIGR>
A;Cross-references: UNIPROT:P24324; GB:U32720; GB:I42023; NID:gl573322; PIDN:AA22013.1;
R;Experimental source: strain Rd Kw20
R;Maskell, D.J.; Szabo, M.J.; Butler, P.D.; Williams, A.E.; Moxon, E.R. Mol. Microbiol. 5, 1013-1022, 1991
A;Title: Molecular analysis of a complex locus from Haemophilus influenzae involved in P
A;Reference number: S15287; MUID:92065797; PMID:1956282
A;Accession: S15287
A;Molecule type: DNA
A;Residues: 'MNGTRICOSINOSINOSINOSINOSKSVIIAGNGTSLKSIDYLLPKDYDVFR' ,7,'QYFPE
A;Cross-references: EMBL:X57315; NID:g43587; PIDN:CAA40567.1; PID:g43588
A;Experimental source: strain RM7004
C;Superfamily: Haemophilus influenzae hypothetical protein HI0352
Query Match 28.1%; Score 441.5; DB 2; Length 231;
Best Local Similarity 41.5%; Pred. No. 2.3e-23;
Matches 95; Conservative 46; Mismatches 79; Indels 9; Gaps 6;
QY 67 KLIQEENYNIENIVCST-INLEYIDGQFVDNFELYSDFAPLGHGHEIIKKLDKDFAYIKYN 125
Db 2 OLKKNNEYADILLSFVNLGSE-LKKTKNKKLTQVDIGHYILNKLPAFDAYLQYN 60
QY 126 EYNQRITSGVMYCATAVAGYKSIYISGIDFYQDTNNLYAFDNKKNNLKNCTGFKNQ 185
Db 61 ELYENKRITSGVMYCATAVAGYKSIYISGIDFYQDTNNLYAFDNKKNNLKNCTGFKNQ 120
QY 186 KPFKFNHSMACDLOALDYLKRDVNNIYSLNS-----EYFKLAP-DIGSDFVLSKPKKY 240
Db 121 KQSDIHSMEYDLNLYFLQKHGVNIYCI SPSPCLNCFPLSPLNPNITFILEEK-KNY 179
QY 241 INDILIPDKVAQERY-YGKSRLENLHYKLIKDILRPSDIKHLYKEK 288
Db 180 TQDILIPDKFYKGIYKIPRIYQNLIFRLIWDILRPNIDIKHALKSR 228
RESULT 3
H90124
hypothetical protein orf753 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H90124
R;Douglass, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: H90124
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-753 <DOU>
A;Cross-references: UNIPROT:Q98S90; GB:AF081031; NID:gl3794316; PIDN:AAK39693.1; GSPDB:G
C;Genetics:
A;Gene: orf753
A;Map position: 3
A;Genome: nucleomorph
C;Keywords: nucleomorph
Query Match 9.1%; Score 143; DB 2; Length 753;
Best Local Similarity 23.0%; Pred. No. 0.025;
Matches 67; Conservative 48; Mismatches 102; Indels 74; Gaps 14;
QY 34 NQYFEDRYFVGKDVKYVFN-----PVVFEEQYTSKKLQNEYNENIENI----- 79
Db 478 NKIIFSEYFDNKKFYFYNQNLPHINVQDFIVYKIRKNNLIPYQDKFEVKKPSYLL 537

QY 80 -----VCSTINLEYIDGQFVDNFELYSDF-----AFLGHEIIKKLDKDFAYIKYNEI 127
Db 538 TYNFLAYIKICEMSKI-CIDFKFKINWEIINSEILFOFAGEFCSKKL-NFYIIMKFNPNP 595
QY 128 YNQRITSGVMYCATAVAGYKSIYI---SGIDFY-----QDTNNLYAFDNKK 172
Db 596 LFTSLFNKNFSLSEKELNFPFKIDIIIFSKLNFPSLLIKFFNLHKKNNLYMLNKNL 655
QY 173 KNLLNKTGFKNQKPFKFNHSMACDLOALDYLKRY---DVNIYSLNS---DEYFKLAPDI 227
Db 656 K-IYKLTTPFLNKKFKTN-----LMVKYSLDDDDHIYSCDMLNRLIKGTISI 701
QY 228 GSDFVL--SKPKKY-----INDILIPDKVAQERYYKSKSRLENLHYKLI 271
Db 702 QADFLILNLRKDRYFYFPEVRGIL-----NQIRILNEKLLLKKNVKNYKFL 747
RESULT 4
D90094
hypothetical protein orf471 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D90094
R;Douglass, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: D90094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <DOU>
A;Cross-references: UNIPROT:Q98RR3; GB:AF165818; NID:gl3794509; PIDN:AAK39884.1; GSPDB:G
C;Genetics:
A;Gene: orf471
A;Map position: 1
A;Genome: nucleomorph
C;Keywords: nucleomorph
Query Match 8.2%; Score 129.5; DB 2; Length 471;
Best Local Similarity 24.5%; Pred. No. 0.12;
Matches 62; Conservative 33; Mismatches 85; Indels 73; Gaps 12;
QY 36 FYFEDRYFVGKDVKYVFNPFVFEQYTSKKLQNEEYN-----IENIVCSITNLE 87
Db 170 FKTKMYFLIRIVRY---NDFFFF-----LVKSRFNPNPKILFIYIFFFSLANLS 218
QY 88 YIDGQFVDNFELYSDFAPLGHGHEIIKKLDKDFAYIKYNEIYNNRQITSGVMYCATAVAG 147
Db 219 FPKGMVFPKDPFKFSSSI---KNILSYFDAPFFFLNPFEEKKQKNNQS-----IN 265
QY 148 YK-----SIVISGIDFYQDTNNLYAFDNKKNNLKNCTGFKNQKPK-----FIN 191
Db 266 YQKKKHAIILENFSFQNYCETLLSYDNKSSNNQNCILIFKSKFEKLSLQIYPLNFIN 325
QY 192 HSNACDLOALDYLKRYDVNIYSLNSDEYFKLAPDIGSDPVLK-----KPKKYINDILI 246
Db 326 LN-----DYFMK-----NIYSNKK-----KISNILLSTFKIKKSHFEIKKKKHIN----- 365
QY 247 PKYAOERYYGKK 259
Db 366 ITKLALSKYFFKK 378
RESULT 5
C71610
probable membrane associated protein PFB0615c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: C71610
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;

Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: C71610
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2013
A;Cross-references: UNIPROT:O96216; GB:AB001406; GB:AB001362; NID:G3845230; PIDN:AAC7191
A;Experimental source: clone 307
C;Genetics:
A;Gene: PFB0615C

Query Match 8.2%; Score 128.5; DB 2; Length 2013;
Best Local Similarity 18.9%; Pred. No. 0.75; Mismatches 59; Indels 127; Gaps 15;
Matches 66; Conservative 61; Mismatches 121; Indels 82; Gaps 16;

QY 34 NQVFEDRYFVGDKVYVFPNPFVFE--QVYTSKKLQNEEYNIEN-----IVCSTINL 86
DB 1147 NKTFFNDNMKEDKLEKINEDFVITEGKSKNKIKNTQNDNNNDVFCNSL-- 1204
QY 87 EYIDGFDVDFNPFYSDAFGLGHEIIKGLDFFAYIKYNEI--YNRQRI----- 133
DB 1205 -----YEL-----LLNKEKSFLLNKHGKLYINERMHTSELTYDITV 1243
QY 134 TSGVYMCATAVALGYK-----SIYISGIDFYQDTNNLY-APD 169
DB 1244 TNNILICISFNSVDYPLEINPHINIRYMPYLANNIDIQYYPILIKGNNYNNNNMYDLFL 1303
QY 170 NKKKLNKCTGFGNQ-----KPFNFHSMACDQALD-----YLMKRYDVNIY--SL 215
DB 1304 IKKNFLLRNKKEDEALIKQEKDHSICPKLIONQDQYNTKCVENFVNTI 1363
QY 216 NSDE-----YFKLAPDYGSDPVLSPKPKKYINDI 244
DB 1364 NSNEHISFYLSKWIENNNTSYINDSLIKNNIVFLKIKNDISQNYT-NRKRKNPFEDI 1422
QY 245 LIPDKYAQERYYKSKSLKENL-----HYKLIKDLRLPSDIK 282
DB 1423 VCMEKKYIEN--NKNNEKKNIKVDINNNMPTNYLNKILLNDVE 1470

RESULT 6
S72284
DNA-directed RNA polymerase (EC 2.7.7.6) beta'-2 chain - Plasmodium falciparum plastid
N;Alternate names: RNA polymerase rpoC2
C;Species: plastid Plasmodium falciparum
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 09-Jul-2004
R;Wilson, R.J.M.; Denny, P.W.; Preiser, P.R.; Rangachari, K.; Roberts, K.; Roy, A.; Whyt
J. Mol. Biol. 261, 155-172, 1996
A;Title: Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium fa
A;Reference number: S72277; MUID:96346169; PMID:8757284
A;Accession: S72284
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-960 <WIL>
A;Cross-references: UNIPROT:Q25802; EMBL:X95275; NID:g1171583; PIDN:CAA64574.1; PID:g117
A;Note: biosynthesis of this protein involves a -1 frameshift in the codon for residue 5
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
C;Genetics:
A;Gene: rpoC2
A;Genome: plastid
A;Note: this apparently degenerate plastid is referred to as the apicoplast
C;Superfamily: DNA-directed RNA polymerase beta chain
C;Keywords: nucleotidyltransferase; plastid; transcription

Query Match 8.0%; Score 126; DB 2; Length 960;
Best Local Similarity 21.4%; Pred. No. 0.48;
Matches 72; Conservative 61; Mismatches 121; Indels 82; Gaps 16;

QY 4 NINALCGGSPSLKNDYKRLPKQFDVFCNQCFEEDRYF--VGK-----DVKY 50
DB 443 NIQLLI-----KNHNKWI--LYNIYTYLYIYHIKFNLYNKGILLNNNNKNYVIY 493

QY 51 VFPNPFVPEQYYSKLLQNEEYNIEN-----IVCSTINLEYIDGFDVDFNDEL 100
DB 494 FLINYNLFSNYYK---YNNNNYFINSNGYFCKGNFILKNFNNTQILNKLFPVNNIFI 550
QY 101 YPSDA-----FLGHEIIKGLDFFAYIKYNEIYNRORITSGVVMCATAVAGYKSI- 151
DB 551 YKYKELFLYINININIIKKILNFKY--TNKGLFFIKKNNNFLYIEIFKYNWYKYL 609
QY 152 -----YISGIDFYQDTNNLYAFDN---NKKLLNKKCTGFKQKPKFNFHNSMA 195
DB 610 LNNKNLYIYNNYIKYKYNININLYFIKNLFYNNNNFIHNNIYKNNYIY--NNNN 668
QY 196 CDLQALDYLMKRYDVNIYSLNDEYFKLAPDYGSDPVLSPKPKKYINDILIPDKYAQERY 255
DB 669 -----LYQVKNKILNINNNLNNKLF---YNNINNNIYLNLDITIGLQSNIIIF 715
QY 256 YGKSKSLKENLHVKLIKDLRLPSDIKHYLKEKYAN 291
DB 716 ENK--NIKONIFF--ISNNIYFIFYIKY---NYLN 744

RESULT 7
A44396
P-type cation translocating ATPase - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jun-2000
C;Accession: A44396
R;Kriegha, S.; Cowan, G.; Meade, J.C.; Wells, R.A.; Stringer, J.R.; Robson, K.J.
J. Cell Biol. 120, 385-398, 1993
A;Title: A family of cation ATPase-like molecules from Plasmodium falciparum.
A;Reference number: A44396; MUID:93132070; PMID:8421054
A;Accession: A44396
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1984 <KRI>
A;Note: sequence extracted from NCBI backbone (NCBI:P122678)

Query Match 8.0%; Score 125; DB 2; Length 1984;
Best Local Similarity 22.8%; Pred. No. 1.3;
Matches 81; Conservative 42; Mismatches 90; Indels 142; Gaps 20;

QY 24 LPKQFDVFCNQFYEDRYFVGVKDYVFFNFVFEQYYSKLLIQ----- 70
DB 1436 LSKYKYVHHKYYYPDSCITNLRKKNSLFYNLKKYI--YEEKKYLQHLKLDHNNYKVE 1493
QY 71 -----NEEYNIENIVCSTINLEYIDGFDVDFNPFYSDAFGLGHEIIKGLDFFAYIK 123
DB 1494 LPRIKDIYNSYQWESI--KTRNFIHLSQFA-----FSNLLSYIILN--DDNNVYK 1544
QY 124 ---YNE--IYNRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLY----- 166
DB 1545 NYIYKNNYIYNNKNSICNNYICNNYIYNNKNIY-----NKNYIYNNKNIILTHAKSV 1596
QY 167 -----APDN-----NKKN-----LNKCT-GPK--- 193
DB 1597 LLSGSSKKFLKPFNSIIRHKLKKNKKNIKRYKNNHVNNTSKGHILNMTGHPKDY 1656
QY 184 ---NOKFKFNFHSMACDQALDYLMKR---YDVNIYSLNDEYFKLAPDYGSDPVLSPK 237
DB 1657 SSLKKNYIRVNNKR-----YMLKNDVYDRHMYNL--TDMY-----RGTYGCSKKK 1701
QY 238 KKYI-----NDILIPDKYAQERYYKSKSLKENLHYKLI-----KDLRLPSDIKH 283
DB 1702 NKNIYNNNNNII-----KNNKINRFLHLLVDKCKRNICHYTIDKN 1743

RESULT 8
H71606
hypothetical protein PFB0755w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: H71606

G97169
Protein containing a domain related to multimeric flavodoxin wrba family [imported] - C1
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97169
R;Nolling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, R.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97169
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-729 <KUR>
A;Cross-references: UNIPROT:Q97H25; GB:AB001437; PIDN:AAK80146.1; PID:gl5025184; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2188

Query Match 7.7%; Score 121.5; DB 2; Length 729;
Best Local Similarity 21.1%; Pred. No. 0.71;
Matches 58; Conservative 45; Mismatches 75; Indels 97; Gaps 14;
QY 26 KQDFVRCNQFYRED-----RYFVGKDVXVF-----FNPVFEEQYTSKKLIQ 70
DB 49 KPHVLNVMDVFEDDEENRFFLSNKKELFILTLYKLYNRVFSNFYIES-----S 100
QY 71 NEETNIE-----NIVCTSTINLEY-IDGFOFVDNFELFSDAFGLGHEIIEKKLDFPAYIK 123
DB 101 EARTNVQLDEYLNLLINK-KIDYKISKNFITSNENPFIDDDICG-ENVKKLED----- 152
QY 124 YNEINRQRTSGVYMCATVALGYKSIYISGID-----FYQDTNNLYAFD--NNKQNL 176
DB 153 -----TKELIEAKEDKQFYRYFKDKYMFEPYDGEKNII 188
QY 177 NKCTGPKNQKFK-----FINHSMACD--LQALDVLKRYDVNIYSLNSDEYVEKLA PDIGS 229
DB 189 RLDNGSKNEKINYNTIFINSNPFALDFEALGNISK--STKCPMNEYAFAFKLATD--- 243
QY 230 DFVLSKKPKYINDILIPDKYQRYGKKSRLKE 264
DB 244 -----IKDEYFKLKHFKQKFFYIKQ 262

RESULT 12
G71607
probable integral membrane protein PFB0735c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: G71607
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: G71607
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1182 <GAR>
A;Cross-references: UNIPROT:O96240; GB:AB001414; GB:AB001362; NID:g3845260; PIDN:AAC7193
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0735c
Query Match 7.7%; Score 121.5; DB 2; Length 1182;
Best Local Similarity 23.1%; Pred. No. 1.2;
Matches 58; Conservative 34; Mismatches 84; Indels 75; Gaps 14;
QY 56 FVFEQYTSK--KLQNEEYNIENIVCTINLEYIDGFOFVDNFELFSDAFGLGHEIIEK 113
DB 819 FTIIQKFDSSMIQIQNDKNNYQHISCHNYSKP-----KDNSEYIPD---DHN--- 865
QY 114 KLKDFPAYIKYNEINRQRTSGVYMCATVALGYKSIYISGIDFYQ-DTNNLYAFDNNK 172

Db 866 --KLLVNY-SYNQLYEKNHND-----DNIFHDLKIYERNINNKYQKIKDK 909
QY 173 KXLLNCKTGFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFV 232
DB 910 K-----KIYAFKN-KINLINIPLCN-----NVKEHFSFNPVANNIKY 946
QY 233 LSKKPKK-----YINDILIPDKYQRYGKKSRLKENLHYKLIKDLIRLPSDIKH----- 283
DB 947 QRTPENISKLMYINN-----SQBFQNTQKONFFPHILNYSLYTHIKNNP--IKKQNTN 997
QY 284 --YLKEKYANK 292
DB 998 NLYIKNDYNNQ 1008

RESULT 13
D90103
hypothetical protein orf419 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D90103
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reit
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: D90103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-419 <DOU>
A;Cross-references: UNIPROT:Q9XG38; GB:AJ010592; NID:g4583666; PIDN:CAB40404.1; GSPDB:GN
C;Genetics:
A;Map position: 2
A;Gene: orf419
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 7.6%; Score 119.5; DB 2; Length 419;
Best Local Similarity 21.2%; Pred. No. 0.52;
Matches 72; Conservative 58; Mismatches 99; Indels 111; Gaps 17;
QY 19 IDYKRLPKQFDV-----FRNQYFVEDRY-----FYGKDV 48
DB 119 INLKNLEHNFDFLMHKSKITLFSHSFERFLVNFKPFVLFMDRYMIYDQSVTFGLKIL 178
QY 49 KY-----VFNPFPVFQYTSKKLIQNEEYNIENIVC-STINLEYIDG-----FQ 93
DB 179 NHSNIISFN--LIFKNIFLSKKL---KKYSFEGVLLQNDINLKFIDTHVCILSCNLLYN 233
QY 94 FVDNPELYFSDAPLGHIEIKKLDKDFAYIKYNIYNRQRTSGVYMCATVALGYKSIYI 153
DB 234 SIKNFYFF-----GSLILMLYFNPFRSHSYSENF-KKKISNVIL----- 272
QY 154 SGIDFQDTNNLYAFDNNKKNLLKCTGPKNQKFKFINHSMACDLQALDYLKRYDVNI- 212
DB 273 -----FCQ-----IYVPER-----FSIKSFICRNQOQKTLNQTNYTCLFNEINLN 313
QY 213 ----YSLNSDEYFKL--APDIGSDFVLSKKPK--KYINDILIPDKYQRYGKKSRLK 263
DB 314 NVTAHSENIGEYLYIYEIKNFYKRYLVFKQKINIKLNTIWFKIRYDQTLPKKYRSN 373
QY 264 E---NLHYKLIKDLIRLPS-----DINKYLKEKYANK 292
DB 374 EPLNHLIYFLTKNVEYLYTFFNSINILKIKNHQNKQFRNK 413

RESULT 14
T28180
hypothetical protein ORF19 - Melanoplus sanguinipes entomopoxvirus (strain Tuscon)
C;Species: Melanoplus sanguinipes entomopoxvirus
A;Variety: isolate Tuscon
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C;Accession: T28180
R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A;Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A;Reference number: 220484; MUID:99102612; PMID:9847359
A;Accession: T28180
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-437 <AFO>
A;Cross-references: UNIPROT:Q9YW73; EMBL:AF063866; NID:g4049647; PIDN:AA097851.1; PID:94
A;Experimental source: isolate Tuscon
C;Genetics:
A;Note: MSV019

Query Match 7.5%; Score 118; DB 2; Length 437;
Best Local Similarity 23.2%; Pred. No. 0.69;
Matches 71; Conservative 46; Mismatches 93; Indels 96; Gaps 17;
QY 48 VKVFFNPFVFFQYTSKLIQNEEYNIENIVCSTINLE-----YIDG-----F 92
DB 48 LKNIFYKDLIVKGGIASQLITQIDNYNIRGDLDISINIEPNKFNKIILDNIKNKNKTY 107
QY 93 QFVDPNPELYSDAFELGHEIIKKLDF---FAYIKYNEIYNRQITSGVYMCATAVALGYK 149
DB 108 KITDIFE-YINDEKLINDISQIMENYKIEFEQISDNIDFNMMVPISSKYNNNIYIAGAYN 166
QY 150 SIY--ISGIDFYODTNLNYAFDNNKXLLNKCTGFKNOKKFINH---SMACDLQALDY 203
DB 167 KYNVNISNIKYGK-----FNNLKFVN-----IVDNKILIRFYTYIKFACNLK---- 211
QY 204 LMKRYD-----VNISYLSNS---DEYFKLAPD-----IGS 229
DB 212 --KHDTFVIEKDKYCKKFLPLDISITPLKTNINNIEINTFLDNKYNVINQIOTLPH 269
QY 230 DFVLSKPKKYIND-ILIPKIAQE-----RYGKK---SRLK-----ENL---HYK 269
DB 270 SFLKHKIKKRVNRFIILFNKYKNELTQNSQNRRYIDNKKILYSRIKFIENDENILYNY 329
QY 270 LIXDLI 275
DB 330 FFKDLI 335

RESULT 15
H97269
Zn-dependent peptidase, insulinase family [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence.revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97269
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H97269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-976 <KUR>
A;Cross-references: UNIPROT:Q97EVO; GB:AE001437; PIDN:AAK80947.1; PID:g15026063; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3006

Query Match 7.4%; Score 117; DB 2; Length 976;
Best Local Similarity 21.9%; Pred. No. 2;
Matches 64; Conservative 51; Mismatches 105; Indels 72; Gaps 17;
QY 18 NIDYKGLPKQDFVR-----CNQYFEDRYFV---GKOVKYYFFN---PFVFFEQYTS 65
DB 714 NLD-ENFSEKFEKVKLKGICDTIFNKDNVIVSVTGNDEVYKFKELPKIKLKSFEVP 772
QY 66 KKLQNEEY-----NIENIVCSTINLEY-IDGFQFVDNFELYFSDAFELGHEIIKKLDF 118

Db 773 AK----SEYKNKVAQSNNEGLITSS-KVQYAAKGFN-----RRLGYDYSGRMKVL 818
QY 119 FAYIKYNEIYNRQITSGVYMCATAVALGYKSIYISGIDFYQDNTNLYAFDNNKKNLANK 178
DB 819 KSIISLSYLWNNVRVWVGAYGCSAYIQRNGSILFAS-----YRDPN-----LTE 862
QY 179 CTCGKNQKPKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKK 238
DB 863 TLKIYDEVYKYE-----NFNADDYEMTKYILG--TISSIDQPLLPKQIG-----QKSDS 910
QY 239 KYINDILIPD--KYAQRYYGKKRUKENLHYKLIKOLIR-----LPSDIK 282
DB 911 YFENKLTIEDLQKERDEILSTKKEIKS--YSKLLKDVMEQNYICVLGNDVK 960

Search completed: June 13, 2005, 21:10:09
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 16:47:14 ; Search time 4914 Seconds
(without alignments)
2899.030 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNIALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKVR 294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10735419/runat_13062005_165724_12648/app_query.fasta_1.455
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10735419 @CGN 1.1 4200 @runat_13062005_165724_12648 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1571	100.0	3071	1 AF195055	Campylobacter jejuni
2	1571	100.0	4668	1 AF257460	Campylobacter jejuni
3	1571	100.0	13484	1 AF400047	Campylobacter jejuni
4	1571	100.0	13484	1 AY044156	Campylobacter jejuni

5	1571	100.0	282183	1	CJ11168X4	Campylobacter jejuni
6	823	52.4	876	6	AF216647	Campylobacter jejuni
7	823	52.4	876	6	BD249791	Campylobacter jejuni
8	823	52.4	876	6	AR271700	Sequence
9	823	52.4	876	6	AR481782	Sequence
10	823	52.4	876	6	AR527381	Sequence
11	823	52.4	876	6	AX934425	Sequence
12	823	52.4	11474	1	AF130984	Campylobacter jejuni
13	823	52.4	11474	6	BD249790	Campylobacter jejuni
14	823	52.4	11474	6	AR271699	Sequence
15	823	52.4	11474	6	AR481781	Sequence
16	823	52.4	11474	6	AR527380	Sequence
17	823	52.4	11474	6	AX934424	Sequence
18	823	52.4	12576	1	AY297047	Campylobacter jejuni
19	817	52.0	876	6	BD249793	Campylobacter jejuni
20	817	52.0	876	6	AR271702	Sequence
21	817	52.0	876	6	AR481784	Sequence
22	817	52.0	876	6	AR527383	Sequence
23	817	52.0	876	6	AX934429	Sequence
24	816	51.9	11455	1	AY044868	Campylobacter jejuni
25	813	51.8	876	6	BD249792	Campylobacter jejuni
26	813	51.8	876	6	AR271701	Sequence
27	813	51.8	876	6	AR481783	Sequence
28	813	51.8	876	6	AR527382	Sequence
29	813	51.8	876	6	AX934427	Sequence
30	813	51.8	12388	1	AF401529	Campylobacter jejuni
31	812	51.7	11442	1	AF400048	Campylobacter jejuni
32	804	51.2	873	6	AR481786	Sequence
33	804	51.2	873	6	AR481787	Sequence
34	804	51.2	873	6	AR527385	Sequence
35	804	51.2	873	6	AR527386	Sequence
36	804	51.2	873	6	AX934434	Sequence
37	804	51.2	873	6	AX934436	Sequence
38	804	51.2	876	6	BD249794	Campylobacter jejuni
39	804	51.2	876	6	AR271703	Sequence
40	804	51.2	876	6	AR481785	Sequence
41	804	51.2	876	6	AR527384	Sequence
42	804	51.2	876	6	AX934431	Sequence
43	804	51.2	4749	1	AF305571	Campylobacter jejuni
44	804	51.2	11474	1	AF215659	Campylobacter jejuni
45	804	51.2	12390	1	AF401528	Campylobacter jejuni

ALIGNMENTS

RESULT 1	AF195055	Campylobacter jejuni strain MSC57360	DNA	linear	BCT 06-NOV-2000
LOCUS	AF195055	(s1ac), and s1aa (s1aa) genes, complete cds.			
DEFINITION	AF195055				
ACCESSION	AF195055				
VERSION	AF195055.1	GI:11095586			
KEYWORDS					
SOURCE		Campylobacter jejuni			
ORGANISM		Campylobacter jejuni			
REFERENCE	1	(bases 1 to 3071)			
AUTHORS		Guerry, P., Ewing, C.P., Moran, A.P. and Trust, T.J.			
TITLE		Distinct Functional Alleles of Sialic Acid Biosynthetic Genes in			
JOURNAL		Campylobacter Species			
REFERENCE	2	(bases 1 to 3071)			
AUTHORS		Guerry, P., Ewing, C.P., Moran, A.P. and Trust, T.J.			
TITLE		Direct Submission			
JOURNAL		Submitted (14-OCT-1999) Enteric Diseases Dept., Naval Medical			
FEATURES		Research Center, 8901 Wisconsin Ave., Bethesda, MD 20889, USA			
source		Location/Qualifiers			
		1. .3071			
		/organism="Campylobacter jejuni"			
		/mol_type="genomic DNA"			
		/strain="MSC57360"			
		/db_xref="taxon:197"			

TITLE Direct Submission
JOURNAL Submitted (18-APR-2000) Enteric Diseases, Naval Medical Research Center, 503 Robert Grant Ave., Silver Spring, MD 20910, USA
FEATURES Location/Qualifiers
source 1. .4668
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="MSC 57360"
/serotype="HS01"
/db_xref="taxon:197"
gene 53. .937
/notes="type strain of HS01 serotype"
CDS 53. .937
/gene="cstIII"
/genes="cstIII"
/notes="Cst; involved in sialylation of LOS core"
/codon_start=1
/transl_table=11
/product="2,3-sialyl transferase"
/protein_id="AAG37020.1"
/db_xref="GI:11527838"
/translation="MSNNALVCGNGPSLKNIDYKRLPKQPDVFRNCFYEDRYFV
GDKVYVFNPFVFFQYTSKLIQNEEYNIENIVCSTINLEYIDGFOFVDFNFELYP
SDAPLGHIEIKKLDPAFYIKYNEIYNRORITSGVMCAFAVALGYSIVISGIDFYO
DTNLYAFDNKNLNLKCTGFKNOKEKFNHSMACDLOALDYLKRYDVNIYSLNSD
EYKLPADIGSDFVLSKPKYINDILIPDKYQERYGKSKRLKENLHYKLIKOLIR
LPSDIKHYLKBYANKNR"
918. .1949
/gene="neuB1"
918. .1949
/gene="neuB1"
/notes="neuB1"
/codon_start=1
/evidence="not experimental"
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAG37021.1"
/db_xref="GI:11527839"
/translation="MQIKDKLTISQKNPLIIPFINGININGSLEIAKLMVDAAKAGA
KIHKQTHIVEDSMSEQKQNVIPGNANI SIYEIMQCALNYKDEALKEVVEKQGLVY
LSTPFGRAANRDMGSAVKIGSGECNNYPLIKHIAQFKKPMIISTGMSIESIKP
TVKILRDYETPFVLLHTNLYPTPSHLVRLQAMLELYKGFNCLYGLSDHTNNLACIG
AIALGASVLERHTESMDKGPDI VCSMDESLKDLINQTEMVLLRGDNKNPLKEB
QVTIDFAPASVSIKDIKGEILLMDNIWVRFSKGGISAKDFEALIGRRAKQDIIKN
IQUTWDDFE"
1946. .3061
/gene="neuC1"
1946. .3061
/gene="neuC1"
/notes="neuC1"
/codon_start=1
/transl_table=11
/product="N-acetyl mannoseamine synthesis protein"
/protein_id="AAG37022.1"
/db_xref="GI:11527840"
/translation="MKKIVFVSQTRADFSKIKSLMKVNSNEFELPIFVFGHMSKK
FGSTYMEIKCGPKNYKINHDKYYQMDKALSSITDGFSGFKTHEIEPLIVVHGDRV
EPLAAIVGSLNNILVAHIEGSGELSGTIDESLRAISKLAIHLVNDIEAKRLIQMG
EDEKSIPIGSPDLNNNTISLDEAKKYVDIKFNVAIAIPIPI TEINSLFYKQSEB
FVNALIKSEKNYIVIPNNDLGFELILQNVRLKNNRPKIPPSLFYEFYISLLKNAN
FLIGNSCIIKEALYNILGVSQDGRDINKTIRVNAEKDILEAILNTSKTN
ITNKRLEILNLSSEQFRLUKNNILFTINKQKIFMDKK"
3058. .4668
/gene="cgt/neuA1 fusion"
3058. .4668
/gene="cgt/neuA1 fusion"
/note="Cgt/NeuA1; apparent fusion of beta
1,4-N-acetylglucosaminyltransferase and CMP-NeuNAc
synthetase"
/codon_start=1
/transl_table=11
/product="beta 1,4-N-acetylglucosaminyltransferase/
CMP-NeuNAc synthetase fusion protein"

/protein_id="AAG37023.1"
/db_xref="GI:11527841"
/translation="MTLFYKIIAFRLRLKIDKKLPDNEYFLNLNKKIYNEKHGKPFDP
FPNSKDTKSPINPWPFAIRVQNEATLRVLSLEMLPAIQRGVIGYNDCTDGSSEIILE
FKQYSPSFVVKYPHEVQIENQSEENKLSHYNNYVASFIPQDEMIKIDCNVSFIERM
KLVQSPYMAKNTAVRPFRIFLIDKIVIONIGECGFIIDGGDOLLIKCNVSFIERM
VSKQSQMIDPEKTVKELYSEQOILPKHILQAELOMHFPALKYHRNDYQKHLDAIT
LEDFKIHYHRKIKKINTYMLDEKVIIRILDKFKLSGKMTLAIIPARAGSGIKNKK
NLALLHRLPLYTTINAAKNSKYVDKIVLSSDGDIDILEYGTQGVVDVLRKPKSELADD
TTSDKVLVLTLYFYKDYENIVLLQPTSPLRNTNVIDEFLKFNKNSNLISVVECDN
KILKAPIDIDNGNLKIGICDNKYPFMPQKLPKTYMSGAIYIVKSNLFLANNPTFLQSKT
SCVIMDEKASLDIDTTEDLKRVNNISFL"
ORIGIN
Alignment Scores:
Pred. No.: 7,17e-116 Length: 4668
Score: 1571.00 Matches: 294
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-735-419-10 (1-294) x AF257460 (1-4668)
Qy 1 MetSerMetAsnIleAsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAsp 20
Db 53 ATGAGTATGAATTAATTCCTTTGGTATCGGTAATGACCTAGTTTAAAAAATATAGAT 112
Qy 21 TyrLysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAsp 40
Db 113 TATAAAGGTACCAAAACAATTTGATGATTATAGATGCAATCAGTTTATTTTGAAGAT 172
Qy 41 ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPhePheGlu 60
Db 173 AGATATTTTGTGGTAAAGATGTAAATATATGTTTTTAACTCCTTTGTGTTTTTGAA 232
Qy 61 GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleVal 80
Db 233 CAATACACTACTAGTAAAAAATCTATACAAAACGAAAGATATATAACATAGAAAATATAGT 292
Qy 81 CysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeu 100
Db 293 TGTCTACTATTAAATTTAGAATACATAGATGGATTTCAATTTTGTGATTAATTTTGAATTA 352
Qy 101 TyrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPhePheAla 120
Db 353 TACTTTAGTGATGCTTTTTTAGTGCATGAGATTATATAAAAGCTTAAAGATTTTTTTGCC 412
Qy 121 TyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCys 140
Db 413 TATATTAAATATAAGTAAATATACAAATGCGCAAGAAATTAAGTGGAGTTTATATGTGT 472
Qy 141 AlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGln 160
Db 473 GCTACAGCTGTGGCTCTGGGATATAAAGTATATATATAAGTGGTATTGATTTTATCAA 532
Qy 161 AspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysLeuLeuAsnLysCysThr 180
Db 533 GATACAAATAATTTATATCCCTTCGATAATAATAAAAAAAATTTATATAAATAAATGCACT 592
Qy 181 GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
Db 593 GGTTTTAAAAAATCAAAAAATTTCAATTTTCAATTCATTCGATGCGATGTTTACAGCT 652
Qy 201 LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr 220
Db 653 TTAGATTATCTAATGAAAAGGTATGATGGAATATTATATAGTTAAATTCAGATGAGTAT 712
Qy 221 PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
Db 713 TTTAAATTCGCTCCGGATATTGGAAGTGATTTTGTATTGATTAAGTAAAAAATATAT 772
Qy 241 IleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer 260


```
ORIGIN
LLRLPSDIKHFKGK"

Alignment Scores:
Pred. No.: 6.22e-57 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 6 Gaps: 6

US-10-735-419-10 (1-294) x AF216647 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLeuAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATTATTGCTGCAAAATGACCAAGTTTAAAGAAATGATTATTCAGAGACTACCAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
QY 48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAGGCGAGTATTTACAAATCTTCTTTTGAACAAATACACTACTCTTTAAACAT 192
QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAATCCAAAATCAAGATATGAGCCGAACATAATATGTTGTTCTTAATACACCAAGCT 252
QY 88 TyrlleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAAATGAAAATTTTGAAAACTTTTACGATTTATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTCAGTACGATAGCCCTAGGA 432
QY 148 TyrlleSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTCCGGAATGATTTTATCAAAATCGGTCACTCT--TATGCT 489
QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACTAAACAAAAATCTTTTAAATTCGCTCCTTAATTTTAAATATGATAATTCA 549
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAATACTATATTCCTTATGTCCTCAACAGCTCTTTTACAAATTTTATAGAACTA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysTyrIleAsnAsp 243
Db 670 GCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db 775 AATTTTAAAAATAAATAAAGAAAAATTTATTTATCAAGTTGATAAAGATCTATTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTGATATAAGCAATTATTTCAAGGAAAA 873
```

```
RESULT 7
LOCUS BD249791 876 bp DNA linear PAT 17-JUL-2003
DEFINITION Campylobacter glycolyltransferases for biosynthesis of gangliosides and ganglioside mimics.
ACCESSION BD249791
VERSION BD249791.1 GI:33059561
KEYWORDS JP 2002535992-A/2.
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycolyltransferases for biosynthesis of gangliosides and ganglioside mimics
JOURNAL Patent: JP 2002535992-A 2 29-OCT-2002; NATIONAL RESEARCH COUNCIL OF CANADA
COMMENT OS Campylobacter jejuni
PN JP 2002535992-A/2
PD 29-OCT-2002
PF 01-FEB-2000 JP 2000597438
PR 01-FEB-1999 US 60/118213,31-JAN-2000 US 06/495406 PI MICHEL GILBERT, WARREN W WAKARCHUK
PC C12N15/09,C12N1/21,C12N9/10,C12N9/86,C12N15/00 CC Campylobacter glycolyltransferases for biosynthesis of CC gangliosides and ganglioside mimics
CC ganglioside mimics
FH Key Location/Qualifiers
FT source 1..876
FT /organism='Campylobacter jejuni'.
FEATURES
source
1..876
/organism='Campylobacter jejuni'
/mol_type='genomic DNA'
/db_xref='taxon:197'
ORIGIN
Alignment Scores:
Pred. No.: 6.22e-57 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 6 Gaps: 6
US-10-735-419-10 (1-294) x BD249791 (1-876)
QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATTATTGCTGCAAAATGACCAAGTTTAAAGAAATGATTATTCAGAGACTACCAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
QY 48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAGGCGAGTATTTACAAATCTTCTTTTGAACAAATACACTACTCTTTAAACAT 192
QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAATCCAAAATCAAGATATGAGCCGAACATAATATGTTGTTCTTAATACACCAAGCT 252
QY 88 TyrlleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAAATGAAAATTTTGAAAACTTTTACGATTTATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
```



```

Db      670  GCGCCAAATTAATTCATATTTATCATACAGAAAAA---ATAACTACATAAGAT 726
Qy      244  ILleuileProAspLysTyrAlaGlnGluArgTyrGly----- 257
Db      727  ATACTCATACCTTCTAGTGAGGCT-----TATGCAAAATTTTCAAAAAATATT 774
Qy      258  -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLeuLysAspLeuLeu 275
Db      775  AATTTTAAAAAATAAAATAAAGAAAAATTTTATTTACAAAGTTGATAAAAGATCTATTA 834
Qy      276  ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
Db      835  AGATTACCTAGTATATAAGCAATATTTCAAAGGAAA 873

RESULT 11
AX934425
LOCUS      AX934425
DEFINITION Sequence 2 from Patent WO02074942.
ACCESSION AX934425
VERSION    AX934425.1 GI:40641672
KEYWORDS
SOURCE
ORGANISM   Campylobacter jejuni
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Campylobacteraceae; Campylobacter.

REFERENCE
1
Gilbert,M. and Wakarchuk,W.W.
Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
Patent: WO 02074942-A.2 26-SEP-2002;
National Research Council of Canada (CA)
Location/Qualifiers
1..876
/organism="Campylobacter jejuni"
/mol_type="unassigned DNA"
/db_xref="taxon:197"

1..876
/notes="unnamed protein product; bifunctional
alpha-2,3/alpha 2,8-sialyltransferase Campylobacter
sialyltransferase II (CstII) from C. jejuni strain OH4384
(ORF 7a of lipooligosaccharide (LOS) biosynthesis locus)"
/codon_start=1
/transl_table=11
/protein_id="CAF04166.1"
/db_xref="GI:40641673"

CDS
1..876
/translations="MKVVIAGNGPSLKEIDYSLRPNDFVFRCNQFYFEDKYLGGK
CKAVFNPILFFQYVTLKHLIONQVETELMCSNVNOAHLNENFVKTFDYPPDA
HLGYDFKQLKFNAPKPEIYFNQRIITSGVTMCVAIALGYKEIYLSGIDPYQNGS
SYAFDTKQKLLKLPFNKDNNSHIIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLGA
NFTELAPNLSNFIIOEKNNYTKDILIPSEAYGKFSKNINFKKIKIRENIYYKLIKD
LLRLPSDIKHYPKKG"

ORIGIN
Alignment Scores:
Pred. No.:      6.22e-57      Length:      876
Score:          823.00      Matches:    161
Percent Similarity: 69.62%      Conservative: 43
Best Local Similarity: 54.95%      Mismatches: 71
Query Match:    52.39%      Indels:    18
DB:             6           Gaps:      6

US-10-735-419-10 (1-294) x AX934425 (1-876)

Qy      8  LeuValCysGlyAsnGlyProSerLeuLysAsnLeuAspTyrLysArgLeuProLysGln 27
Db      13  ATATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAGACTACCAATGAT 72
Qy      28  PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73  TTTGATGATTAGATGAATCAATTTTATTTTGAAGATAAATACTATCTTGTGTAATA 132
Qy      48  ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrSerLysLys 67

```

```

Db      133  TGCAGGCGAGTATTTTACAAATCTTATCTCTTTTGTGAAACATACACTTAAACAT 192
Qy      68  LeuileGluAsnGluGlyTyrAsnileGluAsnileValCysSerThrileAsnLeuGlu 87
Db      193  TTAATCCAAATCAAGATATAGAGCGCAACTAATTTATGTGTCTTAATTACAACCAAGCT 252
Qy      88  TyrileAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253  CATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCTGTGTCATTTG 312
Qy      108  GlyHisGluileileLysLysLeuLysAspPheAlaTyrileLysTyrAsnGluile 127
Db      313  GGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAATTTTCAGAAAT 372
Qy      128  TyrAsnArgGlnArgileThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373  TATTTCAATCAAGAAATTAACCTCAGGGTTTATATGTGTGACGATAGCCATAGCA 432
Qy      148  TyrLysSerileTyrileSerGlyileAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433  TACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATCGGTCATCT---TATGCT 489
Qy      168  PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490  TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCTTAATTTTAAATATGATTAATCA 549
Qy      188  LysPheileAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550  CACTATATCGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208  TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610  TACAAATATAAACTATATTGCTTATGTCCTCAACAGCTTTTACCAATTTTATAGAACTA 669
Qy      224  AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrileAsnAsp 243
Db      670  GGCCCAATTTAAATTCAAATTTTATCATACAGAAAAA---ATAACTACACTAAGAT 726
Qy      244  IleLeuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727  ATACTCATACCTTCTAGTGAGGCT-----TATGCAAAATTTTCAAAAAATATT 774
Qy      258  -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuileLysAspLeuile 275
Db      775  AATTTTAAAAAATAAAATAAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTA 834
Qy      276  ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
Db      835  AGATTACCTAGTATATAAGCAATTTTTCAAAGGAAA 873

RESULT 12
AF130984
LOCUS      AF130984
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,
            partial sequence.
ACCESSION AF130984
VERSION    AF130984.1 GI:6940827
KEYWORDS
SOURCE
ORGANISM   Campylobacter jejuni
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Campylobacteraceae; Campylobacter.
            1 (bases 1 to 11474)
            Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J.,
            Cunningham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.
            Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
            Identification of the glycosyltransferase genes, enzymatic
            synthesis of model compounds, and characterization of nanomole
            amounts by 600-mhz (1h) and (13)c NMR analysis
            J. Biol. Chem. 275 (6), 3896-3906 (2000)
            20127862
            PUBMED
            10660542

```



```
QY 244 IleuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 6774 ATACTCATACCTTCTAGTGGCT-----TATGAAATTTTCAAAAAATATT 6821
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuileLysAspLeuile 275
Db 6822 AATTTTAAAAAATAAAATTAAGAAATATTATTACAGTTGATAAAGATCTATTA 6881
QY 276 ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
Db 6882 AGATTACCTAGTATAAAGCATTTATTTCAAGGAAAA 6920

RESULT 15
AR481781
LOCUS AR481781 11474 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6699705.
ACCESSION AR481781
VERSION AR481781.1 GI:47243416
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert M. and Wakarchuk W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 1 02-MAR-2004;
FEATURES
source
1. 11474
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 8.35e-56 Length: 11474
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 6 Gaps: 6

US-10-735-419-10 (1-294) x AR481781 (1-11474)

QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 6060 ATTATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTTCAAGACTTACCAAAATGAT 6119
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 6120 TTTGATGATTATAGATGAATCAATTTTATTTGAAGATTAATACTATCTTGGTAAAAA 6179
QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 6180 TGCAAGCGCAGTATTTTACATCTCTATCTTTTGAACAATACTACACTTTAAACAT 6239
QY 68 LeuileGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 6240 TTAATCCAAATCAAGAAATATGAGACCGAAGTAAATTTGTTCTTAATTACAAACCAAGCT 6299
QY 88 TyrileaspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 6300 CATCTAGAAATGAAATTTTGAAGAACTTTTACGATTATTTCTGATGCTCATTTG 6359
QY 108 GlyHisGluileLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluile 127
Db 6360 GGATATGATTTTTCACCAACTTAAGATTTTAAGTCTTATTTAAATTTACGAAAT 6419
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 6420 TATTTCAATCAAGAAATTAACCTCAGGGCTTATATGTGTGATGCTAGCCCTAGGA 6479
QY 148 TyrLysSerileTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
```

Search completed: June 14, 2005, 20:11:30
Job time : 5010 secs

```
Db 6480 TACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 6536
QY 168 PheAspAsnAsnLysLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 6537 TTTGATACTAAACAAAAAATCTTTTAAAAATGGCTCCTAAATTTTAAAAATGATAATTCA 6596
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 6597 CACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAACT 6656
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 6657 TACAAATTAACACTATATTGCTTATGCTCCTACAGCTCTTTACCAAAATTTTATAGAACTA 6716
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 6717 GCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 6773
QY 244 IleLeuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 6774 ATACTCATACCTTCTAGTGGCT-----TATGGAATAATTTTCAAAAAATATT 6821
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuileLysAspLeuile 275
Db 6822 AATTTTAAAAAATAAAATTAAGAAATATTATTTATCAAGTTGATAAAGATCTATTA 6881
QY 276 ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
Db 6882 AGATTACCTAGTATATAAAGCATTTATTTCAAGGAAAA 6920
```

This Page Blank (uspto)

XX 01-FEB-1999; 99US-0118213P.
 PR 31-JAN-2000; 2000US-00495406.
 XX (CANA) NAT RES COUNCIL CANADA.
 XX Gilbert M, Wakarchuk WW;
 XX WPI; 2000-524418/47.
 DR P-PSDB; AAY97204.
 XX Novel glycosyltransferase polypeptides and polynucleotides useful for
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
 PT reagents and as immunogen for producing antibodies.
 XX Claim 6; Page 90; 120pp; English.
 PS
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
 CC useful for synthesizing sialylated oligosaccharides such as ganglioside,
 CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing
 CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyze the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems.
 CC (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.2e-73 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 Gaps: 6

US-10-735-419-10 (1-294) x AAA53721 (1-876)

Qy 8 LeuValCysGlyAnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
 Db 13 ATTATTGCTGGAAATGGCCAGCTTTTAAAGAAATTTGATTATTCAGACATACCAAAATGAT 72
 Qy 28 PheAspValPheArgCysAnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 73 TTTGATGATTATAGATTGATCAATCTTTTATTTGAAGATAAATACATCTCTGGTAAATAA 132
 Qy 48 ValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSerLysLys 67
 Db 133 TGCAAGGCAGTATTTTACAATCCTATCTTTTGTGAAACAATACATACACTTTAAACAT 192
 Qy 68 LeuIleGlnAsnGluTyrAnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 Db 193 TTAATCCAAATCAAGAAATGAGACCGCAACTAATATTATGTGTTCTTAATACCAACCAAGCT 252
 Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 253 CATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATATTCTCTGATGCTCATTTG 312

108 GlyHisGluIleIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle 127
 Db 313 GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTTCAGAAATT 372
 Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 Db 373 TATTTCAATCAAGAATTACCTCAGGGGTTTATATGTGTGCAGTAGCCATGAGCCCTAGGA 432
 Qy 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla 167
 Db 433 TACAAAGAATTTATCTTTCCGGAATTTATTTTATCAAAATGGGCACTCT---TATGCT 489
 Qy 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 Db 490 TTTGATACTAAACAAAAAATCTTTTAAATTCGCTCTCTAAATTTTAAAAATGATAATTCA 549
 Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 Db 550 CACTATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTAGAAAAAACT 609
 Qy 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 Db 610 TACAAATAAAACTATATTGCTTATGTCTCTAACAGTCTTTTAGCAAAATTTTATAGAACTA 669
 Qy 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 Db 670 GCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
 Qy 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 Db 727 ATACTCATACCTCTTAGTGAGGCT-----TATGGAAAAATTTTCAAAAAATATT 774
 Qy 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
 Db 775 AATTTTAAAAAATAAATAATTAAGAAAAATTTATTACAGTTGTATAAAGATCTATTATTA 834
 Qy 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
 Db 835 AGATTACCTAGTATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 2
 ABT13666
 ID ABT13666 standard; DNA; 876 BP.
 XX
 AC ABT13666;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE C. jejuni bifunctional sialtransferase catII coding sequence #1.
 XX
 KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 XX WO200274942-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-CA0000229.
 XX
 PR 21-MAR-2001; 2001US-00816028.
 XX
 PA (CANA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 XX WPI; 2003-040554/03.
 DR P-PSDB; ABJ18479.

DR P-PSDB; AAY97211.
XX Novel glycosyltransferase polypeptides and polynucleotides useful for
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
PT reagents and as immunogen for producing antibodies.
XX
PS Claim 6; Page 94; 120pp; English.
XX
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
CC useful for synthesizing sialylated oligosaccharide such as ganglioside,
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analyzing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyze the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity, compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
XX (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.83e-73 Length: 876
Score: 817.00 Matches: 159
Percent Similarity: 69.62% Conservative: 45
Best Local Similarity: 54.27% Mismatches: 71
Query Match: 52.01% Indels: 18
DB: 3 Gaps: 6

US-10-735-419-10 (1-294) x AAY53725 (1-876)
QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
DB 13 ATATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCACAGACTACCAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
DB 73 TTTGATGATTAGATGCAATCAATTTTATTTTGAAGATAAATACTATCTGTGTAAGAAA 132
QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
DB 133 TGCAGAGCAGTATTTTACAACTCTAGCTCTTTTGTGAACAACTACTTAAACAT 192
QY 68 LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
DB 193 TTAATCCAAATCAAGAAATGAGACCGCACTAATCATGTGTCTTAATTTTACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
DB 253 CATCTAGAAAATCAAAATTTGTAAAAAATTTTACGATTATTTTCTGATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
DB 313 GGATATGATTTTTCACAACTTAAAGAAATCAATGCTTATTTTAAATTTTACGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
DB 373 TATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCACAGTACCCATAGCA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla 167

DB 433 TACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGATCATCT---TATGCT 489
QY 168 PheAspAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
DB 490 TTTGATACCAACCAAAAAATCTTTTAAAAATTTGGCTCTCTTAATTTTAAAAATGATAATTCA 549
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
DB 550 CACTATATCGCATAGTAAATAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
DB 610 TAGCAATAAAGCTATATTGTTTATGCTCCTACAGCTCTTTTAGCAAAATTTTATAGAATA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
DB 670 GCGCCAAATTTAAATTCAAATTTTATCATCAAGAAAAA---AATAACTATATAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly-----TATGGAATAATTTACAAAAATATT 774
DB 727 ATACTCATACCTTCTAGTCAGGCT-----TATGGAATAATTTACAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
DB 775 AATTTTAAAAAATAAATAAATAAAGAAATATTATTACAAGTTGATAAAGATCTATTATA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
DB 835 AGATTACCTAGTATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 6
ABT13668
ID ABT13668 standard; DNA; 876 BP.
XX
AC ABT13668;
XX
DT 07-FEB-2003 (first entry)
XX
DE C. jejuni bifunctional sialtransferase catII coding sequence #3.
XX
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW Galnac transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
XX WO200274942-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA0000229.
XX
PR 21-MAR-2001; 2001US-00816028.
XX
PA (CANA) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
PI
XX WPI; 2003-040554/03.
XX
DR P-PSDB; ABJ18481.
XX
XX New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Claim 8; Page 97; 107pp; English.
XX
XX The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention

CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-Acetylgalactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present DNA sequence represents a Campylobacter jejuni CC gene of the invention

SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.83e-73 Length: 876
 Score: 817.00 Matches: 159
 Percent Similarity: 69.62% Conservative: 45
 Best Local Similarity: 54.27% Mismatches: 71
 Query Match: 52.01% Indels: 18
 DB: Gaps: 6

US-10-735-419-10 (1-294) x ABT13668 (1-876)

QY 8 LeuValCysGlyAanglyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
 DB 13 ATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGACTACCCAAATGAT 72
 QY 28 PheAspValPheArgCysAangInPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 DB 73 TTTGATGATTATAGATGCAATCAATTTTATTTGAAGATATAATACATCTTGTGTAAAAA 132
 QY 48 ValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSerLysLys 67
 DB 133 TGCAGAGCAGTATTTACATCTAGTCTTTTTTTTGAACAATACATCTTTTAAACAT 192
 QY 68 LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 DB 193 TTAATCCAAATCAAGAAATATGAGCGCACTAATCATGTGTCTTAATTTTACCAAGCT 252
 QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 DB 253 CATCTAGAAATCAAAATTTTGTAATAAATCTTTTACGATATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluIleLysLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle 127
 DB 313 GGATATGATTTTTTCAACAACCTTAAAGAAATTCATTTTAAATTTTCAGAAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 DB 373 TATTTCAATCAAGAATATACCTCAGGGGTCTATATGTGCACAGTAGCCATAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
 DB 433 TACAAAGAAATTTATCTTTTCGGGAATGATTTTTATCAAAATGATCATCT--TATGCT 489
 QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 DB 490 TTTGATACCAACAACAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATATGATATTC 549
 QY 188 LysPheIleAsnHisSerMetalCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 DB 550 CACTATATCGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAA 609
 QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 DB 610 TACGAATAAAGCTATATTGTTTATGTCTTAAACAGTCTTTTACCAATTTTATAGAACTA 669
 QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysTyrIleAsnAsp 243
 DB 670 GCGCCAAATTTAAATTCNAATTTTATCATACAGAAAAA---NATAACTATATAAGAT 726

QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 DB 727 ATACTCATACCTCTAGTCAGGCT-----TATGGAAATTTTCAAAAAATATT 774
 QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
 DB 775 AATTTTAAAAAATAAATAATTAAGAAATATATTATTACAGTTGATAAAGATCTATT 834
 QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
 DB 835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 7

AAAS3724
 ID AAAS3724 standard; DNA; 876 BP.
 XX
 AC AAAS3724;
 DT 15-SEP-2003 (revised)
 DT 22-DEC-2000 (first entry)
 XX
 DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
 XX
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside; ds.
 XX
 OS Campylobacter jejuni; O:10 serotype.
 XX
 FH Location/Qualifiers
 FT CDS
 FT 1..876
 FT /*tag= a
 FT /product= "alpha-2,3-sialyltransferase"
 XX
 PN WO200046379-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 01-FEB-2000; 2000WO-CA000086.
 XX
 PR 01-FEB-1999; 99US-0118213P.
 PR 31-JAN-2000; 2000US-00495406.
 XX
 PA (CANA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2000-524418/47.
 DR P-PSDB; AAY97210.
 XX
 PT Novel glycosyltransferase polypeptides and polynucleotides useful for
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
 PT reagents and as immunogen for producing antibodies.
 XX
 PS Claim 6; Page 92; 120pp; English.

XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
 CC useful for synthesizing sialylated oligosaccharide such as ganglioside,
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing

QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAAATAAACAATATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTATAGAAGTA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGGCAATTAATTAATTTATCTACAGAAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
Db 775 AATTTTAAAAAATAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTATATAAGCAATTATTTCAAAGGAAAA 873
RESULT 12
ABT13669
ID ABT13669 standard; DNA; 876 BP.
XX
AC ABT13669;
XX
DT 07-FEB-2003 (first entry)
XX
DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.
XX
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
XX WO200274942-A2.
XX
XX 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA000229.
XX
PR 21-MAR-2001; 2001US-00816028.
XX
XX (CANA) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
PI
PI WPI; 2003-040554/03.
DR
DR P-PSDB; ABJ18482.
XX
XX New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
XX Disclosure; Page 98; 107pp; English.
XX
XX The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present DNA sequence represents a Campylobacter jejuni

CC gene of the invention
XX
SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.8e-71 Length: 876
Score: 804.00 Matches: 158
Percent Similarity: 69.28% Conservative: 45
Best Local Similarity: 53.92% Mismatches: 72
Query Match: 51.18% Indels: 18
DB: 8 Gaps: 6
US-10-735-419-10 (1-294) x ABT13669 (1-876)
QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGGGTACCNAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGTATTAGATGTAATCAATTTTATTGAGATAAATACTATCTTGTGTAATAAAA 132
QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAAAGCAGTGTGTTTACACCCCTAATTTCTTTGAGCAATACTACTTTAAAAACAT 192
QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAAUCCAAATCAAGATATGAGACCGAACTAATATGTTGTTCTTAATTACACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAATGAAAAATTTGTAAAAAATTTTACGATTATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLysAspPheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTTTAAACCAACTTAAAGAAATTAATGCTTATTTTAAATTTTCAAGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAATTAACCTCAGGGTCTATATGTGTGCGATAGCCATAGCCCTTAGA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAAGAAATTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCACT---TATGCT 489
QY 168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACCAACCAAGAAAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAATGATCGCTCG 549
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGGACATAGTAAATAATCAGATATAAAGCTTTTAGAATTTCTAGAAAAAAT 609
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAAATAAACAATATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTATAGAAGTA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
Db 775 AATTTTAAAAAATAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTATATAAGCAATTATTTCAAAGGAAAA 873

PD 03-APR-2003.
 XX 26-SEP-2002; 2002WO-JP009907.
 XX 26-SEP-2001; 2001JP-00292796.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Endo T, Koizumi S;
 XX WPI; 2003-393339/37.
 DR P-PSDB; ABR40195.
 XX Alpha 2,3/alpha 2,8 sialyltransferase from Pasteurella multocida for
 PT production of sialic acid-containing complex sugars.
 XX Claim 3; Page 59-60; 66pp; Japanese.
 XX The present invention relates to a method for producing alpha-2,3/alpha-
 CC 2,8-sialyltransferase. alpha-2,3/alpha-2,8-sialyltransferase can be used
 CC in a method for producing sialic acid-containing complex sugars from an
 CC oligosaccharide substrate using a culture of a microorganism transformed
 CC with DNA encoding alpha-2,3/alpha-2,8-sialyltransferase, or an extract of
 CC the culture. The present sequence is a coding sequence for alpha-
 CC 2,3/alpha-2,8-sialyltransferase from Pasteurella multocida
 XX
 SQ Sequence 909 BP; 305 A; 131 C; 154 G; 319 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,59e-58 Length: 909
 Score: 673.50 Matches: 137
 Percent Similarity: 62.80% Conservative: 47
 Best Local Similarity: 46.76% Mismatches: 94
 Query Match: 42.87% Indels: 15
 DB: 8 Gaps: 7
 US-10-735-419-10 (1-294) x ACCT71694 (1-909)
 QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
 DB 40 ATTGTTGCAGGAATGGCGAGAGTTTAAGTCAAAATTTGATTATAGTTGTTACCGAAAT 99
 QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 DB 100 TATGATGTGTTCGTTGTAATCAATTTTATTTGAAGACGCTATTTTATAGGAACAAG 159
 QY 48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
 DB 160 ATAAAGACGATTTCTTCACGCCAGGGGCTTTTCTTGAGCAATATTATACACTTTATCAT 219
 QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 DB 220 CTCAGAGAAACAATGAGTATTTTGTGATAATGTGATCTCTCTCTTTTATCATCCT 279
 QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 DB 280 ACAGTAGAT---TTAGAAAGAGCTCAGAAATACAGCACTTTTATTGATGTGATCAAC 336
 QY 108 GlyHisGlu---IleIleLysLysLysLysAspPhePheAlaTyrIleLysTyrAsnGlu 126
 DB 337 GGATATGAAAAGTATTATCTAACTCAACTCAGCTGCTTTTATTTGGCTATAAAGAA 396
 QY 127 IleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeu 146
 DB 397 TTATATGAGATCAAGAATTACATCTGGGTATATATGTGTGCACTTGTCTATTGGCATG 456
 QY 147 GlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyr 166
 DB 457 GGATATACAGATATTTACTTAACTGATCGATCTGATTTTATCAAGCGGAGCAAGAAACTAC 516
 QY 167 AlaPheAspAsnAsnLysLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLys 186
 DB 517 GCATTCGATATAAAAAGCCCTAATATTATAGGTTATTCGCTGATTTTCGGAAGAAAGAA 576

QY 187 PheLysPheIleAsnHisSerMetalaCysAspLeuGlnAlaLeuAspTyrLeuMetLys 206
 DB 577 ACACCTCTTTTATCATAGATAATGATTGGAAGCATTTATCTTTTACACAG 636
 QY 207 ArgTyrAspValAsnIleTyrSerLeuAsnSer-----AspGluTyrPheLys 222
 DB 637 CATTATCATGTTAAATTTTATTCAATTTTACCAATGAGCCCTTTGTCTAAACATTTTCT 696
 QY 223 LeuAlaProAspIleGlySerAsp-----PheValLeuSerLysLysProLys 238
 DB 697 ATT---CCAACTGTAGAGGATGATTGTGAACACACTTTTGTGGCCACTAATA---GAA 750
 QY 239 LysTyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArg-----Tyr 255
 DB 751 AATTACATTAATGATATATTGTTGCCTCTCTATTGTTGATATGAAAAATTAGGACCATC 810
 QY 256 TyrGlyLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuLys 275
 DB 811 GTGCTCTAAGAAATACGTTTTCATTCTAATTTGTCAGGTTGATTAGAGACTTATTG 870
 QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
 DB 871 AAATTACCGAGTGCCTTAACACTATTTTAAAGAAAA 909
 RESULT 15
 ADT05647/C
 ID ADT05647 standard; DNA; 191996 BP.
 AC ADT05647;
 DT 02-DEC-2004 (first entry)
 XX Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 683.
 DE middle ear bacterial infection; nasopharynx bacterial infection; ds.
 XX Haemophilus influenzae.
 OS WO2004078949-A2.
 PN 16-SEP-2004.
 PD 05-MAR-2004; 2004WO-US007001.
 PF 06-MAR-2003; 2003US-0453134P.
 PR (CHIL-) CHILDRENS HOSPITAL INC.
 PA Bakaletz LO, Munson RS, Dyer DW;
 PI WPI; 2004-662422/64.
 DR New polynucleotides of nontypeable strain of Haemophilus influenzae,
 PT useful for treating or preventing NTHi bacterial infections of the middle
 PT ear and/or nasopharynx.
 XX Claim 1; SEQ ID NO 683; 88pp; English.
 PS The invention comprises nucleotide sequences (genes) from the genome of a
 CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
 CC sequences of the invention are useful for treating or preventing NTHi
 CC bacterial infections of the middle ear and/or nasopharynx. The present
 CC nucleic acid represents an NTHi DNA sequence of the invention.
 XX Sequence 191996 BP; 59302 A; 35787 C; 37096 G; 59811 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.89e-54 Length: 191996
 Score: 663.50 Matches: 136
 Percent Similarity: 64.83% Conservative: 52
 Best Local Similarity: 46.90% Mismatches: 93
 Query Match: 42.23% Indels: 9

```
DB: 13 Gaps: 6
US-10-735-419-10 (1-294) x ADT05647 (1-191996)
Qy 6 AsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuPro 25
Db 43856 TCTGTCATTATGTCAGGTAAATGGAACAAGTTTAAATCAATTGACTATAGTTTATACCT 43797
Qy 26 LysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGly 45
Db 43796 AAAGATTATGATGTTTCCCGTGAATCAATTTATTTTGGAGTCAATTTTCTTGGT 43737
Qy 46 LysAspValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSer 65
Db 43736 AAGAAATATAAAAGGATTTTAAATGTTCTGTAATTTTGAACAATACTATACGTTT 43677
Qy 66 LysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThr---Ile 84
Db 43676 ATGCAATTAATAAAATAATGAATATGATATGCTGATATTTATTTATCATCTTTCTA 43617
Qy 85 AsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAsp 104
Db 43616 AATTAGGGGATTTCAGAA--TTAAGAAATCCAGCGTTTAGAAAAATTACTACCACAA 43560
Qy 105 AlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPheAlaTyrIleLysTyr 124
Db 43559 ATCGATCTTGTCATAGCTATTTAAATAAACTACGAGCTTTTGATGCTCATTTACAAT 43500
Qy 125 AsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaVal 144
Db 43499 CACGAATATATGAGATAAGAGGATTACATCAGCGCTCTATATGTCGAGTGGCACT 43440
Qy 145 AlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsn 164
Db 43439 GCAATGGGTTATAAAGATCTTTATTGACAGCATTCATTTTATCAAGAAAAAGGGAAT 43380
Qy 165 LeuTyrAlaPheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsn 184
Db 43379 CTTTACGCAATTCATCAACAGAAAAATATTATAAATTATTACCTCTCTTTTTCACAA 43320
Qy 185 GlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeu 204
Db 43319 AATATAAGTCAACCGGATATCCATCTATGGAATATGATTTAATGACACTTATTTTAA 43260
Qy 205 MetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAsp-----GluTyr 220
Db 43259 CAAAACATTATGGAGTAATAATTTATTGCAATTCGCCAGAAAGTCCTCTATGTAATTAT 43200
Qy 221 PheLysLeuAlaPro---AspIleGlySerAspPheValLeuSerLysLysProLysLys 239
Db 43199 TTCCCTTTATCACCACCTGAATAACCCCAATTACTTTTATTCTCGAAGAAAG---AAAAAT 43143
Qy 240 TyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyr---TyrGlyLys 258
Db 43142 TACACACAGATATTTTAAATTCGCCGAGTTGTATATAAAAAATTTGGTATATATATCC 43083
Qy 259 LysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuPro 278
Db 43082 AAACCAAGATTACCAAAATCTGATTTTTCGGTTGATCTGGGATATATTACGTTTACCT 43023
Qy 279 SerAspIleLysHisTyrLeuLysGluLys 288
Db 43022 AATGATATAAACACGCTTTGAAAGCAAG 42993
```

Search completed: June 14, 2005, 18:49:07
Job time : 692 secs

Page Blank (uspto)


```

Db      610 TACAAATAAACAATATATGCTTATGCTTACAGCTCTTTTAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProaspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTTTATCATCAACAGAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuLysAspLeuIle 275
Db      775 AATTTTAAAAATAAATAATTAAGAAATATTTATTTACAGTTTGATTAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACTAGTATATAAGCAATTTTCAAGGAAAA 873

RESULT 3
US-10-303-162-2
; Sequence 2, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-2

Alignment Scores:
Pred. No.: 1,1e-92 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-10-303-162-2 (1-876)
Qy      8 LeuValCysGlyAanGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      13 ATATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGACTACCAATGAT 72
Qy      28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGTTTATGATGTAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAA 132
Qy      48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67

```

```

Db      133 TGCAGGCGAGTATTTTACAATCTCTTCTTTTGGAAACAATACACTTTTAAACAT 192
Qy      68 LeuIleGlnAsnGluLysTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      193 TTAATCCAAATCAAGAATATGAGACGAACTAATTATGTTGTTCTTAATTACCAACCAAGCT 252
Qy      88 TyrIleaspGlyPheGlnPheValaspAsnPheGluLeuTyrPheSerAspIlePheLeu 107
Db      253 CATCTAGAAATGAAATTTTGTAAACACTTTTACGATTATTTCTCTGATGCTCATTTG 312
Qy      108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db      313 GGATATGATTTTTCACCAACTTAAGATTTTAAGCTTTATTTTAATTTTCAGAAAT 372
Qy      128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAATTACCTCAGGGGTTTATATGTGTGCAGTAGCATAGCCCTAGGA 432
Qy      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCGGGAATTGATTTTATCAAAATGGGTCTCTCT---TATGCT 489
Qy      168 PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490 TTTGATACTAAACAAAAAATCTTTTAAATTCGCTCTTAATTTTAAATGATAATCA 549
Qy      188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAA 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATAAACAATATATGCTTATGCTTCTTACAGCTCTTTTAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProaspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTTCAAAATTTTATCATCAACAGAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuLysAspLeuIle 275
Db      775 AATTTTAAAAATAAATAATTAAGAAATATTTATTTACAAAGTTGATAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACTAGTATATAAGCAATTTTCAAGGAAAA 873

RESULT 4
US-10-303-134-2
; Sequence 2, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876

```



```
RESULT 13
US-09-495-406-4
; Sequence 4, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; PRIOR FILING DATE: 2000-01-31
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of LOS biosynthesis locus)
US-09-495-406-4

Alignment Scores:
Pred. No.: 1,92e-91 Length: 876
Score: 813.00 Matches: 159
Percent Similarity: 69.97% Conservative: 46
Best Local Similarity: 54.27% Mismatches: 70
Query Match: 51.75% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-09-495-406-4 (1-876)
QY 8 LeuValCysGlyAenGlyProSerLeuLysAenLeuLysAenLeuLysAenLeuLysGln 27
Db 13 ATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGCTTACCAATGAT 72
QY 28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTCTTGGTAAAAA 132
QY 48 ValLysTyrValPhePheAenProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db 133 TTCAAAGCAGTATTTTACAATCTCGTCTTTTGTGAAACAATACTACACTTTAAACAT 192
QY 68 LeuLeuGlnAenGluGluTyrAenLeuGluAenLeuValCysSerThrLeuAenLeuGlu 87
Db 193 TTAATCCAAATCAAGAAATAGAGACCAACTAATATGTTGTTCTAATTAACAACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAATGAAATTTGTAAAAACCTTTTACGATTTATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPheAlaTyrIleLysTyrAenGluLeu 127
Db 313 GGATATGATTTTTTAAACAACAATAAGCAATTTAATGCTTTATTTAAATTTTCAAGAAAT 372
QY 128 TyrAsnAenGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATCTCAATCAAGAAATTAATCTCAGGAGTCTATGTTGTCAGTACTATAGCCCTAGGA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCATCT--TATGCT 489
QY 168 PheAspAenAenLysLysAenLeuLeuAenLysCysThrGlyPheLysAenGlnLysPhe 187

US-09-816-028A-4
; Sequence 4, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: Gilbert, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001110US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-4

Alignment Scores:
Pred. No.: 1,92e-91 Length: 876
Score: 813.00 Matches: 159
Percent Similarity: 69.97% Conservative: 46
Best Local Similarity: 54.27% Mismatches: 70
Query Match: 51.75% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-09-816-028A-4 (1-876)
QY 8 LeuValCysGlyAenGlyProSerLeuLysAenLeuLysAenLeuLysAenLeuLysGln 27
Db 13 ATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGCTTACCAATGAT 72
QY 28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTCTTGGTAAAAA 132
QY 48 ValLysTyrValPhePheAenProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db 133 TTCAAAGCAGTATTTTACAATCTCGTCTTTTGTGAAACAATACTACACTTTAAACAT 192
QY 68 LeuLeuGlnAenGluGluTyrAenLeuGluAenLeuValCysSerThrLeuAenLeuGlu 87
Db 193 TTAATCCAAATCAAGAAATAGAGACCAACTAATATGTTGTTCTAATTAACAACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAATGAAATTTGTAAAAACCTTTTACGATTTATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPheAlaTyrIleLysTyrAenGluLeu 127
Db 313 GGATATGATTTTTTAAACAACAATAAGCAATTTAATGCTTTATTTAAATTTTCAAGAAAT 372
QY 128 TyrAsnAenGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATCTCAATCAAGAAATTAATCTCAGGAGTCTATGTTGTCAGTACTATAGCCCTAGGA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCATCT--TATGCT 489
QY 168 PheAspAenAenLysLysAenLeuLeuAenLysCysThrGlyPheLysAenGlnLysPhe 187
```



```
Db      670 GCGCCAAATTAAATTCAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAAAAATTAAGAAAAATATTATTACAAGTTTGATAAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLysLeuLysGluLys 288
Db      835 AGATTACCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873
```

Search completed: June 14, 2005, 21:15:22
Job time : 227 secs


```

DB: 16 Gaps: 6
US-10-735-419-10 (1-294) x US-10-303-134-2 (1-876)
Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db : : : : :
Qy 13 ATTATTGCTGGAATAGACCAAGTTTAAAGAAATTCATTATTCACAGCTACCAAAATGAT 72
Db : : : : :
Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db : : : : :
Qy 73 TTTGATGATTATTAGATGAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAGAA 132
Qy 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db : : : : :
Qy 133 TGCAGGCGAGTATTTTCAATCCTTATCTTTTGAACAATACTACATCTTAAACAT 192
Qy 68 LeuIleGlnAsnGluGlyTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db : : : : :
Qy 193 TTAATCCAAATCAAGATATGACCGCACTAATTATGTGTTCTAATTACACCAAGCT 252
Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db : : : : :
Qy 253 CATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATATTTTCTGATGCTCATTTG 312
Qy 108 GlyHisGluIleLeuLysLysLysLeuLysLysLeuLysLysLeuLysLysLys 127
Db : : : : :
Qy 313 CGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTTCAGAAAT 372
Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db : : : : :
Qy 148 TyrlsSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db : : : : :
Qy 433 TACAAAGAAATTTATCTTTCGGAATGATTTTATCAAAATGGGTCACTCT--TATGCT 489
Qy 168 PheAspAsnAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 187
Db : : : : :
Qy 490 TTTGATACATAACCAAAAAATCTTTTAAATTTGGCTCCTAAATTTTAAATATGATAATCA 549
Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207

```

RESULT 6

```

US-10-303-162-2
; Sequence 2, Application US/10303162
; Publication No. US20030157658A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

```

```

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)-(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-2

```

```

Alignment Scores:
Pred. No.: 1,57e-74 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 16 Gaps: 6

```

US-10-735-419-10 (1-294) x US-10-303-162-2 (1-876)

```

Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db : : : : :
Qy 13 ATTATTGCTGGAATAGACCAAGTTTAAAGAAATTCATTATTCACAGCTACCAAAATGAT 72
Db : : : : :
Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db : : : : :
Qy 73 TTTGATGATTATTAGATGAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAGAA 132
Qy 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db : : : : :
Qy 133 TGCAGGCGAGTATTTTCAATCCTTATCTTTTGAACAATACTACATCTTAAACAT 192
Qy 68 LeuIleGlnAsnGluGlyTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db : : : : :
Qy 193 TTAATCCAAATCAAGATATGACCGCACTAATTATGTGTTCTAATTACACCAAGCT 252
Qy 88 TyrlsSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 107
Db : : : : :
Qy 253 CATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATATTTTCTGATGCTCATTTG 312
Qy 108 GlyHisGluIleLeuLysLysLysLeuLysLysLeuLysLysLeuLysLysLys 127
Db : : : : :
Qy 313 CGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTTCAGAAAT 372
Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db : : : : :
Qy 373 TATTTCAATCAAGAAATTTACCTCAGGGGTTTATGTGTGCTAGTACCTAGCCCTAGGA 432
Qy 148 TyrlsSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db : : : : :
Qy 433 TACAAAGAAATTTATCTTTCGGAATGATTTTATCAAAATGGGTCACTCT--TATGCT 489
Qy 168 PheAspAsnAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 187
Db : : : : :
Qy 490 TTTGATACATAACCAAAAAATCTTTTAAATTTGGCTCCTAAATTTTAAATATGATAATCA 549
Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207

```

```

Db      550 CACTATATCGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAATAAATAACATATATGCTTATGTCCTTAACAGCTCTTTAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCAAAATTTAAATTTAAATTTATCATCAAGAAAAA---AATAACTACATCAAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTACATCTTCTAGTAGGCT-----TATGAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAATAAATAAAGAAATATTATTACAAAGTTGATAAAGATCTATTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLysLeuLysGluLys 288
Db      835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 7
US-10-735-419-2
; Sequence 2, Application US/10735419
; Publication No. US20040180406A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-735-419-2

Alignment Scores:
Pred. No.: 1-57e-74 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 19 Gaps: 6

US-10-735-419-10 (1-294) x US-10-735-419-2 (1-876)

Qy      8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      13 ATTATTGCTGGAATGGCAAGTCTTAAAGAAATTTGATTATTCAAGACTTACCAATGAT 72
Qy      28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGATTATAGATGTAATCAATTTTATTGTAAGATTAATACTACTCTGTGTAATAAAA 132

```

```

Qy      48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db      133 TGCAGGCGAGTATTTTACATCTCTATCTCTTTTGTGAACATACATACACTTTAAACAT 192
Qy      68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      193 TTAATCCAAATCAAGAATATGAGACCGCAACTAATTATGTGTTCTTAATTACACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAsnAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTATTTCTGATGCTCATTTG 312
Qy      108 GlyHisGluIleIleLysLysLysAspPheAlaTyrIleLysTyrAsnGluIle 127
Db      313 GGTATGATTTTTCACCAACTTAAGATTTTAATGCTTATTTTAATTTTACGAAAT 372
Qy      128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAATTTACCTCAGGGGTTTATATGTGTGAGTAGCCATAGCCCTAGGA 432
Qy      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCATCT---TATGCT 489
Qy      168 PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490 TTTGATACTAAACAAAAAATCTTTTAAAAATGGCTCCTTAATTTTAAAAATGATAATTCA 549
Qy      188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGACATAGTAAAAATAACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATAAATACTATATTGCTTATGCTCTTAACAGCTTTTATAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCAAAATTTAAATTTCAAAATTTTATCATCAAGAAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTACATCTTCTAGTAGGCT-----TATGAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAATAAATAAAGAAATATTATTACAAAGTTGATAAAGATCTATTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLysLeuLysGluLys 288
Db      835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 8
US-10-820-536-2
; Sequence 2, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406

```


us-10-735-419-10.rnpb

Thu Jun 16 13:06:59 2005

373 TATTTCATCAAGAAATACCTCAGGGGTTTATATGTGTGCAGTAGCCATAGCCCTAGGA 432
 148 TrrLysSerIleTrrLysSerGlyLeuAspPheTrrGlnAspThrAsnAsnLeuTrrAla 167
 433 TACAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCTCT---TATGCT 489
 168 PheAspAsnAsnLeuLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 490 TTTGATACCTAAACAAACAAATCTTTTAAATTTGGCTCCTTAATTTTAAATAATGATATTCA 549
 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTrrLeuMetLysArg 207
 550 CACTATATCGGCAATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTAGAAAACACT 609
 208 TrrAspValAsnIleTrrSerLeu-----AsnSer-----AspGluTrrPheLysLeu 223
 610 TACAAATAAAACTATATATGCTTATGCTCAACAGCTTTTAGCAAAATTTTATAGAACTA 669
 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTrrLysAsnAsp 243
 670 GCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTrrAlaGlnGluArgTrrTrrGly----- 257
 727 ATACTCATACCTTCTAGTGAGCT-----TATGGAATAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTrrLysLeuLysLeuLysLeu 275
 775 AATTTTAAATAAAATAAAATAAAAGAAATATTTTATTAACAGTGTATTAAGATCTATTATA 834
 276 ArgLeuProSerAspIleLysHisTrrLysLysGluLys 288
 835 AGATTACCTAGTGTATTAAGCAATTATTTCAAAAGGAAAA 873

RESULT 12

US-10-821-604-2
 ; Sequence 2, Application US/10821604
 ; Publication No. US20040229263A1

GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/821.604
 ; CURRENT FILING DATE: 2004-04-08
 ; PRIOR APPLICATION NUMBER: 10/303,128
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (CSTII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-10-821-604-2

Alignment Scores: 1.57e-74 Length: 876
 Pred. No.:

276 ArgLeuProSerAspIleLysHisTrrLysLysGluLys 288
 835 AGATTACCTAGTGTATTAAGCAATTATTTCAAAAGGAAAA 873
 RESULT 11
 US-10-846-219-2
 ; Sequence 2, Application US/10846219
 ; Publication No. US20040219638A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/846.219
 ; CURRENT FILING DATE: 2004-05-14
 ; PRIOR APPLICATION NUMBER: US/09/816,028
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (CSTII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-10-846-219-2

Alignment Scores: 1.57e-74 Length: 876
 Pred. No.: 823.00 Matches: 161
 Score: 69.62% Conservative: 43
 Percent Similarity: 54.95% Mismatches: 71
 Best Local Similarity: 52.39% Indels: 18
 Query Match: 20 Gaps: 6
 DB:

US-10-735-419-10 (1-294) x US-10-846-219-2 (1-876)
 QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTrrLysArgLeuProLysGln 27
 DB 13 ATTATTCGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAGACTACCAATGAT 72
 QY 28 PheAspValPheArgCysAsnGlnPheTrrPheGluAspArgTrrPheValGlyLysAsp 47
 DB 73 TTTGATGATTTAGATGTTATCAATTTTATTTTGAAGATAAATACTATCTTGTGTAATAA 132
 QY 48 ValLysTrrValPheAsnProPheValPheGluGlnTrrTrrSerLysLys 67
 DB 133 TGCAAGCGATTTTTCACATCTTATTTTTCGACCAATCACTACATCTTAAACAT 192
 QY 68 LeuIleGlnAsnGluTrrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 DB 193 TTAATCCAAATCAAGATATGAGACCAACTAATTTATGTTTCTAATTAACCAAGCT 252
 QY 88 TrrIleAspGlyPheGlnPheValAspAsnPheGluLeuTrrPheSerAspAlaPheLeu 107
 DB 253 CATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluIleLysLysLysAspPheAlaTrrLysTrrAsnGlu 127
 DB 313 GGATATGATTTTTCACACACTTAAAGATTTTAATGCTTATTTTAAATTTTCCAGAAAT 372
 QY 128 TrrAsnArgGlnArgIleThrSerGlyValTrrMetCysAlaThrAlaValAlaLeuGly 147

US-10-735-419-10.rnpb

Thu Jun 16 13:06:59 2005

490 TTGATACAAACAAAAAATCTTTTAAATGGCTCTCTAATTTTAAAAATGATAATTC 549
 188 LysPheLeuSerMetAlaCysAspLeuGlnAlaLeuMetLysArg 207
 550 CACTATATCGGCACATAGTAAAAACACATATAAAGCTTTAGAAATTTCTAGAAAAACT 609
 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 610 TACAAAATAAAACTATATGCTATGCTCTAAGCTCTTTAGCAAAATTTTATAGAACTA 669
 224 AlaProAspIleCysSerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 670 GCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAAAAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
 775 AATTTTAAAAATAAAAAATAAAGAAATATTATTATACAGTTGATAAAGATCTATTA 834
 276 ArgLeuProSerAspIleLysHisTyrLysLeuLysGluLys 288
 835 AGATTACCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 14
 US-10-821-573-2
 ; Sequence 2, Application US/10821573
 ; Publication No. US2004022931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: National Research Council of Canada
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/821,573
 ; PRIOR FILING DATE: 2004-04-08
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 ; OTHER INFORMATION: strain OH384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-10-821-573-2

Alignment Scores:
 Pred. No.: 1,57e-74 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 DB: 20 Gaps: 6

US-10-735-419-10 (1-294) x US-10-821-573-2 (1-876)
 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27

13 ATTATGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTATCAAGACTACCAATGAT 72
 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 73 TTTGATGATTATGATTAATCAATTTTATTTTGAAGATAAATACTACTTCTGGTAAAAA 132
 48 VallysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
 133 TGCAAGGAGGATTTTACAACTCTTCTTTTGTGAAACAATACTACACTTTAAACAT 192
 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 193 TTAATCCAAAATCAAGATATGAGACCGAACTAATATATGTGTCTTAATTAACACCAAGCT 252
 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 253 CATCTAGAAAATGMAAATTTTGTAAAACTTTTACGATTAATTTCTGATGCTCATTTG 312
 108 GlyHisGluIleLeuLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
 313 GGATATGATTTTTCACCAACTTTAAAGATTTTAAATGCTTATTTAAATTTACGAAAT 372
 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 373 TATTTCAATCAAGAAATTTACTCAGGGGTTTATATGTGTGAGTAGCATAGCCCTAGGA 432
 148 TyrLysSerIleTyrIleSerGlyLysAspPheTyrGlnAspThrAsnLeuTyrAla 167
 433 TACAAAGAAATTTTCTTCGGGAATTTGATTTTATCAAAATGGGTCTATCT--TATGCT 489
 168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 490 TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCTCTAATTTTAAAAATGATTAATTC 549
 188 LysPheLeuAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 550 CACTATATCGGCACATAGTAAAAATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 610 TACAAAATAAAACTATATGCTTATGCTCTTAAACAGCTCTTTAGCAAAATTTTATAGAACTA 669
 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 670 GCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAAAAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
 775 AATTTTAAAAATAAAAAATAAAGAAATATTATTATCAAGAAATTTTCAAGGAAAA 873
 835 AGATTACCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 15
 US-10-850-807-2
 ; Sequence 2, Application US/10850807
 ; Publication No. US20040259140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: National Research Council of Canada
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/850,807
 ; CURRENT FILING DATE: 2004-05-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028

; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
 US-10-850-807-2

Alignment Scores:

Pred. No.: 1.57e-74 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Query Match: 54.95% Mismatches: 71
 DB: 52.39% Indels: 18
 Gaps: 6

US-10-735-419-10 (1-294) x US-10-850-807-2 (1-876)

QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
Db	13	ATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGACTACCAATGAT	72
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	73	TTTGATGATTAGATGATCAATTTATTTGAAGATAAATACTATCTTGGTAAAAA	132
QY	48	ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys	67
Db	133	TGCAAGCGCATTTTACAAATCCCTATCTTTTGGAACTACTACACCTTTAAACAT	192
QY	68	LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	193	TTAATCCAAATCAAGATATGAGCCGAACCTAATATGTTCTTAATTACCAACCAAGCT	252
QY	88	TyrIleAspGlyPheGlnPheValAspAsnGluLeuTyrPheSerAspAlaPheLeu	107
Db	253	CATCTAGAAATGAAATTTGTAAACTTTTACGAATTATTTCTCTGATGCTCATTTG	312
QY	108	GlyHisGluIleIleLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127
Db	313	GGATATGATTTTTCAAACCACTTAAAGATTTAATGCTTATTTTAAATTCACGAAAT	372
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
Db	373	TATTTCAATCAAGAAATACCTCAGGGGTTTATGTGTGATGACCAATAGCCCTAGGA	432
QY	148	TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla	167
Db	433	TACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCATCT---TATGCT	489
QY	168	PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe	187
Db	490	TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCTTAATTTTAAATGATAATTC	549
QY	188	LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg	207
Db	550	CATATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAA	609
QY	208	TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu	223
Db	610	TACAAATAAACTATATGCTTATCTCTACAGCTTTTACCAAAATTTATAGAACTA	669

QY	224	AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp	243
Db	670	GCGCCAAATTTAAATTCAAATTTATCATACAAGAAAA---AATAACTACCTAAAGAT	726
QY	244	IleLeuIleProAspLysTyrAlaGlnGluAtgTyrTyrGly	257
Db	727	ATACTCATACCTTCTAGTGAGGCT---TATGGAAAAATTTTCAAAAAATATT	774
QY	258	-----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle	275
Db	775	AATTTTAAAAAATAAAATTAAGAAAAATATTATTATTAACAGTTGATAAAGATCTATTA	834
QY	276	ArgLeuProSerAspIleLysHisTyrLeuLysGluLys	288
Db	835	AGATTACCTAGTATATAAAGCAATTATTTCAAAGGAAAA	873

Search completed: June 14, 2005, 22:46:46
 Job time : 678 secs

Qy 166 TyrAlaPheAspAsnAsnLysLysAsnLeu-----LeuAsn 177
 Db 263 TATATTTTAAATAATAAATAATATAGTGAATATGTTATGATAAAGATACACAT 342
 Qy 178 LysCysThrGlyPheLysAsnGlnLysPheLysPheLeAsnHisSerMetAlaCysAsp 197
 Db 343 ACACATACCAATATATCATCTGAATATATACAAAAAATAACGATACACATAATGAT 402
 Qy 198 LeuGlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnLleTyrSerLeuAsnSer 217
 Db 403 TTTCGTACATG-----TTAATAAATAATACAAATAATATATATAGTTCTTAACAAA 456
 Qy 218 Asp-----GluTyrPheLysLeuAlaProAspLleGlySerAspPheValLeuSerLys 235
 Db 457 GATTACACAATTTAT-----AATGATGAATATTTA----- 486
 Qy 236 LysProLysLysTyrIleAsnAspIle 244
 Db 487 AAAGTGAAGAAATAAATAATGAATA 513
 RESULT 4
 BU497975
 LOCUS PFSTOab91a03.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium
 DEFINITION falciparum 3D7 cDNA 5', mRNA sequence.
 ACCESSION BU497975
 VERSION BU497975.1 GI:22794169
 KEYWORDS EST.
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 1 (bases 1 to 574)
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 AUTHORS Tang,K., Cole,R., Chakrabarti,D., Hayward,R., Clifton,S., Pape,D.,
 Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
 Bowers,I., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
 Tsagarisvili,R., Belyagorod,L., Franklin,C., Carr,L., Grow,A.,
 Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
 Waterston,R., Wilson,R. and Sibley,D.
 WASHU Plasmodium EST Project
 Unpublished (2001)
 CONTACT: L. David Sibley
 WASHU Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library was constructed by Debopam Chakrabarti DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: L. David Sibley
 (sibley@borcim.wustl.edu), Washington University
 Seq primer: -40UP from Gibco
 High quality sequence stop: 421.
 Location/Qualifiers
 1..574
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
 /note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
 XhoI; library was constructed by Debopam Chakrabarti.
 Total RNA samples were isolated from mixed stage
 saponin(0.1%) lysed P. falciparum 3D7 infected
 erythrocytes by the acidic guanidinium-phenol chloroform
 method. The poly A+ RNA was isolated by the polyAT-Tract
 mRNA isolation system (Promega, WI) using streptavidin
 Magnesphere particles. Directional cDNA libraries were
 constructed by oligo d(T) priming of poly(A) + RNA (5mg)
 into EcoRI and XhoI sites of 1 ZapII vector using the Zap
 cDNA synthesis kit (Stratagene, CA). The average size of the
 cDNA inserts in the library was between 1.0 and 1.5kb.
 Clones were mass excised using the ExAssist helper phage

(Stratagene), the phagemids were precitptated with PEG
 8000 and extracted with phenol/chloroform. Phagemid DNA
 was electroporated into DH10B cells."

ORIGIN
 Alignment Scores:
 Pred. No.: 0.00678 Length: 574
 Score: 113.00 Matches: 52
 Percent Similarity: 40.64% Conservative: 37
 Best Local Similarity: 23.74% Mismatches: 76
 Query Match: 7.19% Indels: 54
 DB: 5 Gaps: 11
 US-10-735-419-10 (1-294) x BU497975 (1-574)
 Qy 43 PheValGlyLysAspValLysTyrVal-----PhePheAsnProPheValPhePhe 59
 Db 54 TATATAGATAAAGACACACATTTTATTATTCAAACATTTTATAATAATTTTATTATTAT 113
 Qy 60 GluGlnTyrTyrThrSerLysLysLeuIleGlnAsnGlu-----GluTyrAsnIleGlu 77
 Db 114 ATTGAATAATTATGATAAATATGATCTAATATGCAATATGAACACAGGCAAGCAATATAA 173
 Qy 78 AsnIleValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsn 97
 Db 174 AATATAAAAAATAATAATACGAATGGAACAGAAAAGGAT-----ATTATAAAT 221
 Qy 98 PheGluLeuTyrPheSerAspAlaPheLeuGlyHisGluIleLleLysLysLeuLysAsp 117
 Db 222 GATGAATATTTTATTAATGATGATATGATGATAAAATAAGTTGGAAGTATTGTTATCT 281
 Qy 118 PhePhe-----AlaTyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSer 135
 Db 282 TCTTTCAATGAAATTTATGTAGACATAAATGAATATAT---AAAGACAATATATCACC 338
 Qy 136 GlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGly 155
 Db 339 TCATTGGAA-----AATTATAAGATATGTACATAAAT--- 371
 Qy 156 IleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeu 175
 Db 372 -----AATAATAATAATAGTAATAAT 392
 Qy 176 LeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheLeAsnHisSerMetAla 195
 Db 393 ATGGATGTGATTCATTTATTAATACGTAGAAAAGAAAATATCATACCTTCATTGAT 452
 Qy 196 CysAspLeuGlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeu 215
 Db 453 ---GATAAATATTCATTACAATATTTAAAA----- 479
 Qy 216 AsnSerAspGluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSerLys 235
 Db 480 AATTTTAATGATGATAAATTA-----ATTTTTTAAACAAT 518
 Qy 236 LysProLysLysTyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArg 254
 Db 519 ATGATTTCTAAATATATA-----TATGCAATTAGTATAAACATTTTATTAAGACGG 569
 RESULT 5
 CK573341
 LOCUS 1381 Plasmodium yoelii liver stage LCM cDNA library Plasmodium
 DEFINITION yoelii cDNA, mRNA sequence.
 ACCESSION CK573341
 VERSION CK573341.1 GI:44827431
 KEYWORDS EST.
 SOURCE Plasmodium yoelii
 ORGANISM Plasmodium yoelii
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 894)
 AUTHORS Sacci,J.B. Jr., Ribeiro,J.M.C., Huang,F., Alam,U., Russell,J.A.,
 Blair,P.L., Witney,A., Carucci,D.J., Azad,A.F. and Aguiar,J.C.

FEATURES
 source

Alignment Scores:

```

Pred. No.:      0.0151      Length:      774
Score:          111.50      Matches:      57
Percent Similarity: 36.10%      Conservative: 43
Best Local Similarity: 20.58%      Mismatches: 80
Query Match:      7.10%      Indels:      97
DB:              4          Gaps:      11

US-10-735-419-10 (1-294) x BM167106 (1-774)

Qy 15 SerLeuLysAsnIleAspTyrLysArgLeuProLysGlnPheAspValPheArgCysAsn 34
Db 108 AACATAGGCAAGTTAGCATCAAGAACTAGCAAAACAA---AAGGTAACATCAGTAAT 164
Qy 35 GlnPheTyrPheGluAspArgTyrPheValGlyLysAspValLysTyrValPheAsn 54
Db 165 -----ATATACAAATTTGTTACGACATAGGGCTTACANTGATATATAATTTTAAC 218
Qy 55 ProPheValPhePheGluGlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGlu--- 73
Db 219 ACATATATA-----GAAAAACATGCAGAGATGAAGACAA 254
Qy 74 -----TyrAsnIleGluAsnIleValCysSerThrIleAsn----- 85
Db 255 TTACATTTAGAAATAATTCAACAACAAAAATAATATTTAGTGATAGAAATTAATAATATG 314
Qy 86 -----LeuGluTyr-----IleAsp 90
Db 315 GTTGATTATTAATAAAAAAACCAATTATATACATAATGATGCAAAAGTATTAAATGAT 374
Qy 91 GlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeuGlyHisGlu 110
Db 375 TCGTATAAAAAAAGATAT-----GTATTTTAAAAAACA 413
Qy 111 IleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIleTyrAsnArg 130
Db 414 ATATTAGAAATGTTTGAAGCATATATTTTATTTGTTTGGCATATGAA----- 461
Qy 131 GlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyrLysSer 150
Db 461 ----- 461
Qy 151 IleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPheAspAsn 170
Db 462 -----AATTTATTAACTTTGAAGAA 482
Qy 171 AsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheIle 190
Db 483 AATGATCAGATGTATTAAATATATATATAATATAATATTTTAAACAGATTTAAATTTAAG 542
Qy 191 AsnHisSerMetAla-----CysAspLeuGlnAlaLeu 201
Db 543 GATTATAATGATGATCTCTCTTTTACCTGTTTAAATCTTTAATTTATATATATATGCAATT 602
Qy 202 AspTyrLeuMetLysArgTyrAspValAsn-----IleTyr 213
Db 603 GAAATTTTACAAAAAATTTTATACATATATTTTATTTTAAAAAATGCTGTATAT 662
Qy 214 SerLeuAsnSerAspGluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeu 233
Db 663 AGTATGATATCGAGTTGTTGTTTATATATATATATATATATATATATATATATATATAT 719
Qy 234 SerLysLysProLysTyrIleAsnAspIleLeuIleProAspLysTyr 250
Db 720 AAAAAAAGGGAATAGATATACCAATTTTAAATGATTATTATAGAGCCTAC 770

```

RESULT 7

```

AZ935754      1123 bp      DNA      linear      GSS 21-DEC-2001
LOCUS      AZ935754
DEFINITION      ODG282 Oikopleura dioica Shotgun Library Oikopleura dioica genomic
                  clone G282, genomic survey sequence.
ACCESSION      AZ935754
VERSION        AZ935754.1  GI:17975925

```

KEYWORDS

```

SOURCE      GSS.
ORGANISM    Oikopleura dioica
REFERENCE    Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
              Oikopleuridae; Oikopleura.
AUTHORS      1 (bases 1 to 1123)
              Seo,H.C., Kube,M., Edvardsen,R.B., Jensen,M.F., Beck,A., Spriet,E.,
              Gorsky,G., Thompson,E., Lehrach,H., Reinhardt,R. and Chourrout,D.
              Miniature genome in the marine chordate Oikopleura dioica
              Science 294 (5551), 2506 (2001)
TITLE        Oikopleura dioica
JOURNAL      21625354
MEDLINE      11752568
PUBMED
COMMENT      Contact: Seo HC
              Sars International Centre for Molecular Marine Biology
              Bergen High Technology Centre, Thormohlenstg. 55, N-5008 Bergen,
              Norway
              Tel: 55 58 4389
              Fax: 55 58 4305
              Email: hee-chan.seo@sars.uib.no
              Insert Length: 1123 Std Error: 0.00
              Seq primer: T7 and V4 (ATCCAACGGCTTGGAGACTCT)
              Class: shotgun.

```

FEATURES

```

source
1..1123
/organism="Oikopleura dioica"
/mol_type="genomic DNA"
/db_xref="taxon:34765"
/clone="G282"
/sex="male"
/clone_lib="Oikopleura dioica Shotgun Library"
note="Organ: sperm; Vector: pGEM-T Easy; The shotgun
library was prepared using the EcoRV digested pGEM-T Easy
vector (Promega) by Dr Hee-Chan Seo."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      0.0849      Length:      1123
Score:          107.00      Matches:      68
Percent Similarity: 34.38%      Conservative: 42
Best Local Similarity: 21.25%      Mismatches: 98
Query Match:      6.81%      Indels:      112
DB:              8          Gaps:      15

US-10-735-419-10 (1-294) x AZ935754 (1-1123)

Qy 22 LysArgLeuProLysGlnPheAsp---ValPheArgCysAsnGlnPheTyrPheGluAsp 40
Db 91 AAGCTCATCCCTAAAAATTTGTGACATTATATACAGTCTCCAGAAATTTCTATTTTAATGAA 150
Qy 41 ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPheGlu 60
Db 151 GAA-----TTGAAAAAAGATGGAAACCAATGCCACTTCCACCACCGAAATTTTCAA 204
Qy 61 -----GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGlu 73
Db 205 AATTGTGATGACTCAAGGTCTTCAATATTTTACAGCA---TTTGGAGCGCAAAACAAATTT 261
Qy 74 TyrAsn-----IleGluAsnIle 79
Db 262 TATCGTCAATTTATCTGGCTGTCGATGGATCCAGTACGTCCCGGTTCTAGCTCGATC 321
Qy 80 ValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGlu 99
Db 322 CTTTGTCTCTCTTTTGAAGAAGAGTTTCATCAT----- 354
Qy 100 LeuTyrPheSerAspAlaPheLeuGlyHisGluIleLeuLysLysLeuLysAspPhe 119
Db 355 -----CATGAG----- 360
Qy 120 AlaTyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMet 139
Db 360 ----- 360

```



```
RESULT 9
BJ371051          678 bp  mRNA  linear  EST 08-MAR-2002
LOCUS             BJ371051 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION        BJ371051 Dictyostelium discoideum cDNA clone ddc5p08 5', mRNA sequence.
ACCESSION         BJ371051
VERSION           BJ371051.1  GI:19280434
KEYWORDS          EST.
SOURCE            Dictyostelium discoideum
ORGANISM          Dictyostelium discoideum
REFERENCE         1 (bases 1 to 678)
AUTHORS           Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE             Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL           Unpublished (2002)
COMMENT           Contact: Tadasu Shin-i
                   Center For Genetic Resource Information
                   National Institute of Genetics
                   1111 Yata, Mishima, Shizuoka 411-8540, Japan
                   Tel: 81-559-81-6856
                   Fax: 81-559-81-6855
                   Email: tshini@genes.nig.ac.jp.
FEATURES          Location/Qualifiers
                   1..678
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc5p08"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN
Alignment Scores:
Pred. No.:      0.0502      Length:      678
Score:          106.50      Matches:    59
Percent Similarity: 35.40%  Conservative: 38
Best Local Similarity: 21.53% Mismatches: 80
Query Match:    6.78%      Indels:     97
DB:             4          Gaps:       14
US-10-735-419-10 (1-294) x BJ371051 (1-678)
Qy      46 LysAspValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSer 65
Db      58 AAGAGATATAAAATTAATA-----ATTGAACAATTAAGGGAATTA 96
Qy      66 LysLysLeulle-----GlnAsnGluGluTyrAsn----- 75
Db      97 AAACAATCATTTTAAATATAATATACCGAATAATATGATATGATACNAGTACA 156
Qy      76 -----llegluAsnileVal-----Cys 81
Db      157 ATAACATCAATTTTAAATTAATTAATGAAATGTTTGTAGATAATGGTATTATGCAATC 216
Qy      82 SerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyr 101
Db      217 ACAGAATTGTGATAACAATTAATTTGAAAGTTTTCTTTTACCAAAATTTA----- 267
Qy      102 PheSerAspAlaPheGlyHisGluIleLysLysLysAspPhePheAlaTyr 121
Db      268 TTATCAGAAATCATTTATA-----ACATTAACAACATTTATTCAGATATAATTAAT 318
Qy      122 IleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAla 141
Db      319 AATAATAGTAATAGTAATAATAATAATAATAGTAAA----- 351
Qy      142 ThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAsp 161
Db      352 -----AATAATGAAGAAATTATATTCACAAAAAATTAACCTTTTA----- 390
```

```
162 ThrAsnAsnLeuTyrAlaPhe-----AspAsnAsnLysLysAsnLeuLeuAsnLysCys 179
Db      391 TTTAATATTTTATATAAAATTTATAAATGATAATAATAGTAATAATTTTATCAAGTTAATA 450
Qy      180 ThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGln 199
Db      451 ATTGAATTTCAAAATAGAAAATTTTAAA-----AAACATATAATAATAATTTGATTTAGAA 504
Qy      200 AlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGlu 219
Db      505 ATTCAATGAACAATTT-----TGGAAATTTATAATAAATCAATTT----- 543
Qy      220 TyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLys 239
Db      543 ----- 543
Qy      240 TyrIleAsnAspIleLeulleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLys 259
Db      544 -----TGTTTGATACCAAGATAAATGTAGATCAATTCATTAATTTATATAAAAT 591
Qy      260 SerArgLeuLysGluAsnLeuHisTyrLysLeulleLysAspLeulleArgLeuProSer 279
Db      592 AATAGAATTAGG---AATATTCAAATCACCATT-----TCA 624
Qy      280 AspIleLysHisTyrLeuLysGluLysTyrAlaAsnLysAsn 293
Db      625 TCAGTTAAGATTATATTTTAGAATTCCTTTACCAATAATAAT 666
RESULT 10
BJ365708          833 bp  mRNA  linear  EST 08-MAR-2002
LOCUS             BJ365708 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION        BJ365708 Dictyostelium discoideum cDNA clone ddc36g07 5', mRNA sequence.
ACCESSION         BJ365708
VERSION           BJ365708.1  GI:19275010
KEYWORDS          EST.
SOURCE            Dictyostelium discoideum
ORGANISM          Dictyostelium discoideum
REFERENCE         1 (bases 1 to 833)
AUTHORS           Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE             Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL           Unpublished (2002)
COMMENT           Contact: Tadasu Shin-i
                   Center For Genetic Resource Information
                   National Institute of Genetics
                   1111 Yata, Mishima, Shizuoka 411-8540, Japan
                   Tel: 81-559-81-6856
                   Fax: 81-559-81-6855
                   Email: tshini@genes.nig.ac.jp.
FEATURES          Location/Qualifiers
                   1..833
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc36g07"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN
Alignment Scores:
Pred. No.:      0.114      Length:      833
Score:          104.50      Matches:    62
Percent Similarity: 38.66%  Conservative: 42
Best Local Similarity: 23.05% Mismatches: 88
Query Match:    6.65%      Indels:     77
DB:             4          Gaps:       14
US-10-735-419-10 (1-294) x BJ365708 (1-833)
```



```

Db 471 TATCTGT-----ATCTATCTCTAAATATAATATAATCTATTTATAGA 427
Qy 163 nAsnLeuTyralaPheAsnAsnLysLys-----AnLeuLe 176
Db 426 AAAATGGTATTTTAAACCAATAATAAAAAATGTTATATAAACCAATAGTAA 367
Qy 176 uAsnLysCys---ThrGlyPheLysAsnGlnLysPheLysPheLeAsnHisSerMetAl 195
Db 366 TAAAAATGTTATTTATTCCTCCAAAGGACCAAGATTTTCCACTGCTATATATGATGATC 307
Qy 195 a-----CysAspLeuGlnAlaLeuAspTyLeuMetLysArgTyRAspValAsnIleTy 213
Db 306 ACATATTGTCAATTTGACCATAGAAAAATATTATAAAGATATAATTTATATTCAAT 247
Qy 213 rSerLeuAsnSerAspGluTyPheLysLeuAlaProAspIleGlySerAspPheValLe 233
Db 246 AAATTTAAATTTAAAT----- 229
Qy 233 uSerLysLysProLys-----LysTyRileAsnAspIleLe 245
Db 228 AAATAAATGAATAAATAATTTTAAATAGAAAAATAAATAAAGTAATTAATTAAGTATT 169
Qy 245 uile 246
Db 168 AATA 165

RESULT 12
AZ546559/c
LOCUS
DEFINITION ENTDR06TF Entamoeba histolytica Sheared DNA linear GSS 14-NOV-2000
genomic, genomic survey sequence.
ACCESSION AZ546559
VERSION AZ546559.1 GI:11169417
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 924)
Loftus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 94
High quality sequence stop: 921.
FEATURES
Location/Qualifiers
1..924
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notice="Vector: pHOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for

```

whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

ORIGIN

Alignment Scores: 0.131 Length: 924
Pred. No.: 104.50 Matches: 62
Score: 104.50 Matches: 62
Percent Similarity: 40.46% Conservative: 44
Best Local Similarity: 23.66% Mismatches: 80
Query Match: 6.65% Indels: 76
DB: 8 Gaps: 15

US-10-735-419-10 (1-294) x AZ546559 (1-924)

Qy 16 LeuLysAsnIleAspTyRLeysArgLeuProLysGlnPheAspValPhe----- 31
Db 738 ATTAAATCAATAATAATTAACAAAAAGTGTAAACAAA---TTTAAATATATTTCAAAAAAGTGAAG 682

Qy 32 ArgCysAsnGln-----PheTyRPhelGluAspArgTyRPhelValGlyLysAspValLys 49
Db 681 AGTGCCAAAATATCTTTTATTACTTCAAAAATATTAT-----AATAAGAAA 634

Qy 50 TyRValPhePheAsnPropheValPhePheGluGlnTyTyRThrSerLysLysLeuIle 69
Db 633 AATTATGAAAAATAACGAAGATATAATACTTCAAAAATACGTATGAAATGAAAAATATACG 574

Qy 70 GlnAsnGluGluTyRAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyRile 89
Db 573 AAAAGAGATGAAATAAC---GAAGATATAATCTTCAAAATATC----- 532

Qy 90 AspGlyPheGlnPheValAspAsnPheGluLeu-----TyrPheSerAsp 104
Db 531 -----TTCATTATGAATTTTCGAAATAACGAAGATATAATACTTC----- 490

Qy 105 AlaPheLeuGlyHisGluIleLysLysLeuLysAspPheAlaTyRile---Lys 123
Db 489 -----AAAAAAGAAAAATATCTAAGAAATAGTTCAAAAA 457

Qy 124 TyRAsnGluIleTyRAsnArgGlnArgIleThrSer-GlyValTyRMetCysAlaThrAl 143
Db 456 TATCCAACAATCGTTCATTTTAAATGGAATACCTCAAAAATATCTGTATCTGTTCTTAAAAA 397

Qy 143 aValAlaLeuGlyTyRLeysSerIleTyRileSerGlyIleAspPheTyRLeuAspThrAs 163
Db 396 TATCTGT-----ATCTATTCTAAAAATATAAATAATATCTATTATAGA 352

Qy 163 nAsnLeuTyRAlaPheAspAsnAsnLysLys-----AsnLeuLe 176
Db 351 AAAATGGTATTTTAAAAACAATAATAAAAAATAAAAAATGTTATATAAAAAACAATAGTAA 292

Qy 176 uAsnLysCys---ThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAl 195
Db 291 TAAAAAATGTTATTTATTCCTCCAAAGGACGATTTTCCCTACCTGTGTATATGATGATC 232

Qy 195 a-----CysAspLeuGlnAlaLeuAspTyRLeuMetLysArgTyRAspValAsnIleTy 213
Db 231 ACATATTGTCAATTTGTCATAGAAAAATATTATAAAGATATAATTAATTTCAAT 172

Qy 213 rSerLeuAsnSerAspGluTyRPhelLysLeuAlaProAspIleGlySerAspPheValLe 233
Db 171 AAATTTAAATTTAAAT-----LysTyRileAsnAspIleLe 245

Qy 233 uSerLysLysProLys-----LysTyRileAsnAspIleLe 245
Db 153 AAATAAATGAATAAATAATTTTAAATAGAAAAATAAATAAAGTAATTAATTAAGTATT 94

Qy 245 uile 246
Db 93 AATA 90

RESULT 13
AZ684114/c


```
Db      163  TTAATCT-----TTTGAAACAGTAAAAATAAATGAATATGCTTAATTTTAAACATCGAA 216
Qy      122  IleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAla 141
Db      217  ACATCATATATTAATATATATATACAAACAAATAAAAGATTTAATTAATAATTTATGCT 276
Qy      142  ThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAsp 161
Db      277  ACATCATTTATTAGAAAAATTAATTAATCTTGCAATTGAATCAGGACATGAT----- 324
Qy      162  ThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeuLeuAsn---LysCysThr 180
Db      325  -----AATATAGATGCAGAAATACGAGAAAGAAAGAACAAATTAATGAATTCAAATTTATTA 378
Qy      181  GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
Db      379  AATTTTATAATAATTTTATTATTAATAAAAAATCAAAATATT-----CCTAAAAAGTGT 432
Qy      201  LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr 220
Db      433  ATCGAATATAAGATAAATATATATGATTAATATATATAATAATAATAATTTCCAAATTTTAT 492
Qy      221  PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
Db      493  AATTTTACTGGACTAACAAATAAATTTAAGATTTAAACAATCTCTTAGAGTAAAGTTCAA 552
Qy      241  IleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer 260
Db      553  AATAATAATGAATAAAAAAACTGGAAATTTAT---AATACAAAAAATTTATGGAATAAAAAAT 609
Qy      261  Arg-----LeuLysGluAsnLeuHisTyrLysLeuIleLysAsp 273
Db      610  AATTTATTTTATGAGTTACACAATAATCGGCATGGGGGATTTATTAAGAAT 663
```

Search completed: June 14, 2005, 21:11:37
Job time : 3697 secs

This Page Blank (uspto)